Oslo – A bicycle city for the concerned, the speedy and the rebels

- Exploring difference in preference for infrastructure using a five-dimensional bicyclist typology

Ragnhild Batseba Ødegaard Skaugen

AD NYE VEJE
AALBORG UNIVERSITET

I would like to thank my supervisor Susse Georg and my co-supervisor Aslak Fyhri with the Norwegian Centre for Transport Research for all help and guidance. I would like to thank my beautiful boyfriend Erik for being patient and loving even when I thought I was unlovable. Finally I would like to thank my family for all of their support and constructive discussions.

Ragnhild Skaugen

Oslo 17.09.2015

Abstract

This study explores what bicyclist types we have in Oslo, if these types display different preference for different infrastructure designs and what implications this has for promoting bicycling in Oslo. The bicyclist types were created using a cluster analysis yielding a five-dimensional typology with three types. These were the Conscious Concerned, the Laid-back Rebels and the Speedy Unafraid. An investigation of the bicyclist types' preference for different bicycle infrastructure designs showed that the Conscious Concerned prefer to bike in the immediate environment, but choose the routes with low traffic load; the Speedy Unafraid seek out routes with good flow, separate bicycle paths or large connections, and can go "out of their way" to reach these connections; and the Laid-back Rebel bikes in the immediate environment, but does not necessarily bother to seek out the routes with low traffic load. A denser bicycle network is an essential provision in order to facilitate for bicycling for the first two types, while a bicycle expressway would benefit the Speedy Unafraid.

TABLE OF CONTENT

1 INTRODUCTION

1.1 Rethinking the system of personal transportation

What makes a city is basically concentration of human activities. The activities are dependent on basic goods and services, which are provided through many interlinked systems and structures. The nature of these activities, and the way in which the existing systems are organised, exerts extensive pressure on global ecosystems and local resources (UN-Habitat, 2013; United Nations, 1992). Cities across the world are experiencing problems such as local air pollution, excessive waste generation and overcrowding. Globally, cities contribute greatly to total greenhouse gas emission and hence problems of global warming and climate change (Knox & Pinch, 2014). The pressure experienced is likely to intensify as the urban population is expected to increase from 50 % today to approximately 70 % in 2050 (United Nations, 2014).

This prompts urban planning authorities to rethink the existing structures and systems of the city. They continuously face the choice between proceeding with business as usual, and redesigning the systems to meet the needs of a growing population and to reduce the pressure on global and local resources.

One essential system in every city is the system of personal transport. The system serves a fundamental purpose as it allows its inhabitants access to their everyday needs. It consists of physical infrastructure and services that together provide essential spatial links between social and economic activities in the city (Handy, 2005; Urry, 2007). Today the system is primarily designed to serve private motoring powered by the internal combustion engine (Dennis & Urry, 2009).

The arrangement contributes largely to the pressure exerted on local and global resources. According to the recent IPCC-report about 12 % of the CO_2-emissions caused by energy expenditure can be linked directly to personal transport (Sims et al., 2014). Private motoring demands extensive land-use, it creates local pollutants such as NO_x and dust in suspension, it causes congestion which cost society time and money, it is expensive and it pacifies its users (Dennis & Urry, 2009). These are some of the numerous reasons why activists, academics and politicians around the world voice that the system of personal transport is ripe for rethinking, and that the principal mean for change is to facilitate modal shifts to greener options (Dennis & Urry, 2009; Sims et al., 2014).

In this effort, the bicycle is promoted as a desirable mode of transport, and cities around the world are formulating strategies to increase bicycling in the urban environment. Large cities such as London, New York, Johannesburg and Singapore are actively working and investing in order to increase their bicycling share, and the rationale behind the effort is echoed through the strategies. They emphasise that modal shift from the car to bicycling will reduce congestion, reduce local and global emissions and promote health and equality amongst the city's inhabitants (City of Johannesburg, 2009; Mayor of London, 2015; Singapore Government, 2015; Spacescape, 2014c; The City of New York, 2014).

1.2 Policies to increase bicycling

In order to be able to devise strategies that can successfully promote bicycling it is necessary to understand what factors that influence bicycle use. Due to the current surge in policies to promote bicycling there is a growing interest for the topic, and several studies has been completed aiming to identify new determinants for bicycling (Damant-Sirois, Grimsrud, & El-Geneidy, 2014; Forsyth & Krizek, 2010).

There exists no exhaustive list of factors that can affect the level of bicycling in a city, nor will I try to produce a complete list over the factors proven to affect it. In an anthology from 1992 examining bicycling in urban environments, the author tries to summarise factors assumed to influence bicycle, and successfully illustrates the complexity of the matter (McClintock, 1992). See diagram below.

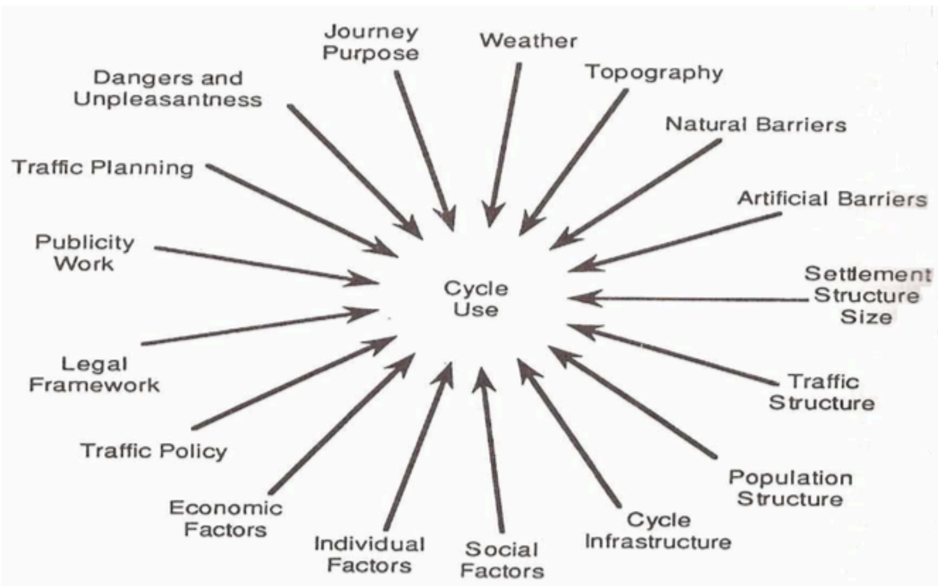


Figure 1 Factors influencing bicycle use (McClintock, 1992, p. 10)

Studies confirm that many of the determinants listed in the figure have an effect on the level of bicycling. A brief selection of proven factors includes weather conditions and elevation (Cleary & McClintock, 2000), social factors such as peer encouragement and the culture and customs in the city (Cleary & McClintock, 2000; Pucher, Dill, & Handy, 2010), the immediate land use pattern (Handy, 2005; Handy, Cao, & Mokhtarian, 2005), the availability and quality of bicycle infrastructure and individual factors such as risk aversion and the level of comfort biking (Hunt & Abraham, 2007; Pucher et al., 2010) and the self-identity as a bicyclist (Handy et al., 2005).

The accumulated knowledge concerning determinants for bicycling has made it possible to formulate appropriate policies and programmes to promote bicycling, and there exists numerous of diverse initiatives. However not all of these are equally effective.

There have been published at least two broad reviews evaluating the relative effect of different types of policy measures to promote bicycling (Forsyth & Krizek, 2010; Pucher et al., 2010). These included a series of soft and physical measures such as programmes and marketing schemes to change behaviour and attitude towards bicycling, education and training, making other modes relatively more expensive, bicycle sharing programmes, bike-to-work days, changing land use patterns and providing new and improving existing bicycle infrastructure (Forsyth & Krizek, 2010; Pucher et al., 2010).

Both reviews concluded that the best strategy to promote bicycling was to combine different types of measures in a comprehensive effort, but that providing bicycling infrastructure needed to be a central part of it (Forsyth & Krizek, 2010; Pucher et al., 2010). Furthermore Forsyth & Krizek concluded that providing bicycle infrastructure is the most important single measure to promote bicycling (Forsyth & Krizek, 2010).

In the following section I will give short introduction of what is actually meant by bicycle infrastructure, and underline the role of bicycle infrastructure in promoting bicycling.

1.4 Introducing bicycle infrastructure

"And which driver is not tempted, merely by the power of the engine, to wipe out the vermin of the street, pedestrians, children and cyclists"

(Theodor Adorno as cited in Dennis & Urry, 2009).

When motorised vehicles became popular around the turn of the twentieth century, it was soon acknowledged that dangerous situations would arise when objects with different velocity moved through the same space (Walker, 2011). Although, hopefully not so premeditated as suggested by culture critic Theodor Adorno in the above comment from 1942. The popular solution to the challenge was to a create zones where the different road users were separated from each other by physical separation, different rights and limitations. The two road users with the highest priority were pedestrians and motorised traffic, and consequently most roadways were separated into two parts, one dedicated to pedestrians and one to motorised transport. This type of demarcation gradually became uniform for the physical roadway structure all around the world, and is one of the reasons why the car dominates the of urban transportation system today. This bisected structure is an essential part of the problem when promoting bicycling today. The bicycle does not have a natural place in the current urban transport system: A bicyclist is faster than the pedestrian, but slower than the car; she is more robust than the pedestrian, but softer than the car (Walker, 2011).

Establishing bicycle infrastructure is in many ways physically claiming a place for the bicycle in the urban transport system where it previously has been marginalised. The aim is that it should be safe and efficient to bike in the city.

There are two main strategies for doing so. One is to continue the philosophy of demarcated system, and creating dedicated zones for bicyclists to minimize interaction. This is called a segregated approach. It is often this strategy that characterises cities known as "bicycle cities", per example Amsterdam and Copenhagen (Forsyth & Krizek, 2010; Walker, 2011). In these cities the bicyclists have own tracks, and do not have to mingle with cars nor pedestrians. The other strategy is to make it more convenient for the bicyclist to mingle, and therefore establish infrastructure that acknowledge the bicyclists place in the modal mix. This is referred to as an integrated approach, and examples of this can be bicycle signs, bicycle parking and bicycle lights (Julien, 2000).

The two approaches to bicycle infrastructure have driven the development of numerous different type of infrastructure designs available for planners and like wanting to promote bicycling in their city. And while the approaches represent to opposing philosophies, many designs and solutions do not necessarily follow this strict polarisation and can be placed on a spectrum ranging from the one extreme to another. I will conclude this section with presenting three broad categories of designs and some of the variations within these categories.

The first category is *bike path*. Here the bicyclists are physically separated from motorised traffic. The level of separation can vary between an elevated track adjacent to the roadway e.g. a cycle track, or paths separated from the road network. The bike paths can be shared with pedestrians or be exclusively for bicyclists. The second category is *mixed traffic*, where bicyclists and other vehicles share the road without any form of separation. This category includes use of traffic calming interventions, shared space solutions and other punctual facilities such as those referred to above. The last category is *bike lane*, which is sort of a "middle child". This is a design where the bicyclists share the road with motorised traffic, but has a separate lane. The lane is indicated with paint in the roadway, often as a line or as a coloured area.

1.5 Manoeuvring the possibilities

The availability of bicycle infrastructure is a necessary and effective measure to increase the share of bicyclists in a city. However, as described in the previous section, there exist several different types of infrastructural designs, each associated with various advantages and disadvantages. Hence, there remains a decision to be made concerning what types of infrastructure is the best fit for a specific location or a specific city.

This is not a straightforward decision to make, as it depends on a series of often conflicting concerns. The capacity of the immediate physical environment, priority between different traffic users, the allocation of resources are examples some of these concerns. A point of departure should however be the type of infrastructure the bicyclists want to use and feel comfortable using. When the ultimate objective is to get more people to bike, improving the bicycle experience seems to be a good place to start.

However all bicyclists are not the same. Some bicyclists have been proven to be more risk averse than others, some evaluate comfortable facilities as the most important factor when bicycling, while others value the possibility to bike fast and make headway without too many stops or barriers (Pucher et al., 2010; Tilahun, Levinson, & Krizek, 2007).

Several studies have shown that dividing bicyclists into groups have yielded more nuanced results than treating them as one homogenous population, and that different interventions are needed to facilitate increased bicycling among various groups of people (Damant-Sirois et al., 2014; Pucher et al., 2010).

This implies that differentiating between different types of cyclists is valuable when planning the provision of bicycle infrastructure. If we want a city that is comfortable to bike in for as many as possible we need to plan for different segments of bicyclists and potential bicyclists.

One of the previously mentioned reviews on the impact of different policy measures, underlines the issue at hand and concludes that the current studies evaluating bicycle infrastructure designs are not sufficient to understand what type of infrastructure is preferred by different bicyclist types (Forsyth & Krizek, 2010).

1.6 Oslo – a bicycle city?

The call for more research described in the previous section coincides perfectly with one of the challenges facing Oslo municipality in the years to come.

The population in Oslo is increasing with 2 % every year, and is one of fastest growing capital in Europe. It is a national objective, affirmed through the Norwegian national transport plan for 2014-2023 that the population growth in the metropolitan areas should not contribute to increased emissions from personal transportation. In order to reach this objective all growth in personal transportation in metropolitan areas should be covered by green transportation modes (Samferdselsdepartementet, 2013). This implies that over the next few years the bicycle share in Oslo needs to increase.

Oslo municipality has risen to the challenge, and have recently developed a new and comprehensive cycling strategy in order to increase the share. The over all vision underpinning this strategy is Oslo as a bicycling city, welcoming all cyclists independent of age and skills (Spacescape, 2014c).

Bicycle infrastructure provision is an important part strategy, and one of the subordinate goals in the strategy to develop a new bicycle network and an Oslo standard for bicycle infrastructure designs (Spacescape, 2014c). Given Oslo’s ambitions to raise the bicycle share by becoming a bicycle city for all, it seems imperative that work with these goals is based on a thorough evaluation of the different types of bicyclists in the city, their needs and preferences.

1.7 Research questions

The purpose of this study is to contribute to the effort of increasing the bicycling share in Oslo and creating a bicycle city that welcomes all types of bicyclists. The above sections illustrate the importance of providing bicycle infrastructure in order to increase the bicycle share, and to make sure that the facilities are suited for a diversified group of bicyclists. This led to the creation of the following research questions:

1. What are the different types of bicyclists in Oslo?
2. How does preference for different infrastructure designs vary between the bicyclist types?
3. What implications do the answers to the above questions have for bicycle infrastructure planning in Oslo?

These questions were investigated quantitatively by employing survey data from bicyclists in Oslo.

1.8 Report structure

The next two chapters will provide the empirical and theoretical context for this study. Chapter 2 describes the conditions for bicycling in Oslo today and present the measures and strategies planned in order to increase the bicycle share. In the third chapter, I will proceed to give a review of both empirical and theoretical literature exploring preference among bicyclists for various bicycle infrastructure designs, and attempt to build a theoretical framework for this study.

Chapter 4 presents the methodology in this study. I will present the survey made for the research project "Safety in Numbers" that provides the data featured in this study and describe how it was adapted to fit the objectives of this study. Following I present the analytical design and the statistical methods employed in this study. The succeeding chapter, chapter 5, describes how the analytical design was made operational e.g. describing the creation of appropriate variables and ensuring that the necessary assumptions for the analytical design are met. In the three subsequent chapters the results from the quantitative tests are presented and analysed. They are organised after each of the research questions. Chapter 7 discusses the relevance of the results. The report is ended with a conclusion summing up the findings and limitations.

2 OSLOVE BICYCLISTS

Oslo - the beautiful and lovable capital of Norway. There are approximately 650 000 people living in Oslo, and every day every one of these makes multiple trips. They make their way to work, to school, to meet friends or to take a hike in forest. Out of all of these trips only 8,3 % is made by bike (Spacescape, 2014b).

The national directions described in the introduction has prompted Oslo municipality to formulate a strategy to increase the bicycle share. In April 2015 they passed a new bicycle strategy for Oslo 2015-2025. Prior to the formulation of this strategy the municipality commissioned a comprehensive evaluation of the conditions for bicycling in Oslo today. This evaluation was normative for the formulation of the new bicycle strategy. The urban consultancy Spacescape was chosen to carry out the evaluation, which included a travel survey, an attitude survey, observations, an evaluation of the previous Oslo bicycle strategy and an evaluation of the existing bicycle network, they also drafted the strategy.

In this chapter I will describe the empirical context of this study. I will describe the conditions for bicycling in Oslo today, including a description of the Oslo bicyclists and an account of state and presence of bicycle infrastructure. In the description of the Oslo bicyclist I heavily rely on the evaluation completed by Spacescape, as this is the dominant source of information on bicycling in Oslo today. Furthermore will describe the public debate about bicycling in Oslo, which often is quite passionate. Finally I will briefly describe the vision and goals underpinning the Oslo bicycle strategy 2015-2025, and elaborate on the elements connected to provision of bicycle infrastructure.

2.1 Bicycling in Oslo today

2.1.1 The Oslo bicyclists

More men than women bicycle in Oslo. The travel survey showed that the genders over all travel the same amount, but when it comes to trips made by bicycle men accounts for a larger share. Bicycling is the mode of travel where the difference between the genders is greatest. The average bicycle trip is 4 kilometres, but men make longer trips than females. When the length of the bicycle trips is combined with the number of trips made, the gender distribution become even more unequal, with women travelling 33 % of all kilometres travelled by bike and men 67 % (Spacescape, 2014b).

The travel study also showed that older adults are more likely to bike than young adults. The bicycle share is highest among those aged 30-59 around 10 % of the trips is made by bike, while young adults aged 18-29 only make 7 % of their trips by bike. The oldest adults, those aged 60-75 is the least likely to bike, with a share of 5 %. No children participated in the travel survey, however the rectors of 86 elementary schools were asked to estimate the number of pupils biking to school in the summer months, and they reported an average share of 15 % (Spacescape, 2014b).

The bicycle share is greatest among those who works (10 %), compared to students (5 %) and retired people (3 %). This coincides with other statistics showing that 12 % of commutes were done by bike, and that this was the trip purpose with the highest share, surpassing leisure purposes (7 %), school (4 %) and service purposes (3 %) (Spacescape, 2014b).

In the attitude survey the respondents were asked about their perception of Oslo as a bicycling city. Only 3 % perceived Oslo as a good city to bike in, and an even lower share, 1 %, agreed that Oslo was a good city to bike in for children and elderly. They were also asked about what the number one initiative to make them start biking or bike more was. 75 % answered that increasing the safety in the bicycle network was the most important initiative. This indicates that the perception of Oslo as a comfortable bicycling city is quite poor, and that many find it unsafe to bicycle in Oslo. Seen in connection with the relatively high level of bicycling in the Oslo this led to the interpretation that there are a lot of "fearless" bicyclists in Oslo, and combined with the gender distribution, this led to the capital statement in the evaluation of the Oslo bicyclist, that the typical Oslo bicyclist is a commuting strong and fearless male (Spacescape, 2014b).

The responses from the attitude survey were used to divide The Oslo population was divided into four profiles based on their responses on the three following questions: (1) Do you usually bike in Oslo in the summer months?, (2) Do you bike no matter what standard the bicycle network hold? and (3) Would you bike more if the bicycle network were safer? (Spacescape, 2014b).

1. *Those who bikes no matter what:* The people in this profile answered yes to the two first questions. They amount to 72 % of those who bike answer that they bike often during the summer months.

2. *Those who only bike where it is safe:* The people in this profile answered yes to the first, but no to the second question. These amount to the remaining 28 % of those say they bike often in the summer months.
3. *The potential bicyclists*: The people in this profile do not bike often in the summer months, but answered yes to the question that they would bike more if the network were safer.
4. *Not interested*: This last profile are those who do not bike often in the summer months, and would not bike more if the bicycle network were safer.

The first two groups constitutes one third of the respondents, the third group another third, and the fourth group the remaining (Spacescape, 2014b). The three first three groups were asked to evaluate different bicycle designs. The results from this exercise are rendered in the coming section about the Oslo standard.

The travel survey by Spacescape, referred to in this section, has been criticized for its methodology. The travel survey was carried out in October. The month was chosen because it best represented the yearly mean share in the years 2004-2012. This based on data from PROSAM. The survey includes respondents between 18 and 75 years (Spacescape, 2014b). This differs from the methodology used in the national travel survey commissioned every fourth year by public agencies[1]. They collect data from the whole year and includes respondents from 13 years and up (Hjorthol, 2014). The results from national travel survey are frequently used for measuring development over time and comparing cities and regions. By not employing a similar methodology in the Spacescape travel survey it becomes isolated and does not relate to previous results yielded from the national travel survey. (Yet, Spacescape concludes in their report that the bicycle share in Oslo has doubled over the past four years, from 4,7 % in 2009 to 8% in 2013 comparing numbers from their survey with the national survey (Spacescape, 2014b). According to the national travel survey the share increased from 4,7 % to 5,4 % (Hjorthol, 2014)). However there is nothing to suggest that the results from the Spacescape survey is “wrong” when the premises for the survey are accounted for.

[1] Avinor, The Ministry of Transport and Communications, The Norwegian Coastal Administration, The Norwegian National Rail Administration and The Norwegian Public Roads Administration.

2.1.2 *The existing bicycle infrastructure*

In Oslo today, per 2015, there is about 180 kilometres roadway that is facilitated for bicycle. For comparison Copenhagen's bicycle network is 412 km. The mask width of the current facilitations, meaning the distance between the roads in the network, is on average 1780 m and 50 % of the inhabitants in Oslo live with 200 metres of bicycle infrastructure (Spacescape, 2014a). These facilities are however not continuous. The pieces and stretches of facilitation are parts of an unfinished main bicycle network. The work with this network started in 1990, when Oslo municipality and the Norwegian Public Roads Administration (NPRA) proposed a joint action plan for establishing a main network for bicycling in Oslo. The network was designed to have its focal point in the city centre. This means that most routes should lead down to the city centre. These are connected through "perpendicular" routes following the existing ring roads (Statens Vegvesen & Oslo Kommune, 2015).

The bicycle facilities is developed complying with the contemporary NPRA's standards for bicycle infrastructure. They operate with four types of infrastructure that qualify for the bicycle network:

1. *Shared pedestrian and bicycle path:* A path physically separated from motorised traffic by a grass border, a ditch, a fence or kerbstones. It is normally only on one side of the road, allowing traffic in both directions. The design is recommended where there is much traffic, high speed limit or if it is outside densely built-up areas.
2. *Bicycle path:* A path dedicated for bicyclists physically separated from motorised traffic by a grass border, a ditch, a fence or kerbstones. It is separated from the pavement using kerbstones or painted markings. The design is recommended where there is much traffic, high speed limit or if it is outside densely built-up areas.
3. *Bicycle lane:* A lane in the roadway dedicated to bicyclists. It is marked with signposts or paint. The design is recommended in densely built areas with medium speed and traffic load.
4. *Mixed traffic:* No specific facilities for bicyclists. The design is recommended in quiet streets with low speed (Statens Vegvesen, 2013).

The last category signifies that stretches with any facilitation can be part of a bicycle network if the street has a low and slow traffic load.

The map below shows the stretches that are facilitated and illustrates the pattern of the network. In addition to the infrastructure designs described above, the map includes recreational trails. These are not an official part of the bicycle network, but can be used by bicyclists provided that pedestrians have the highest priority. The dotted line shows stretches that are part of the bicycle network, but that has yet to be facilitated. These stretches are quite numerous, making the network very fragmented. Another characteristic is how the facility frequently changes from one category to another. Yet, the pattern is that bicycle lanes are the most common type of facility in the inner city, while separated paths are more common further from the centre. The paths are more often than not shared with pedestrians.

Figure 2 Existing bicycle infrastructure in Oslo (Bymiljøetaten, 2015)

In addition to the bicycle facilities described above, bicycles are allowed on all roadways that are no categorised as a motorway or highway, in the public transport lane and on the pavement. Bicycling on the pavement is allowed under the condition that the pedestrian traffic is limited ("Forskrift om kjørende og gående trafikk (trafikkregler)," 1986). Similar to pedestrians, bicyclists are allowed to travel in both directions on the pavement. The arrangement allowing bicyclists on all pavements can only be found in Norway and Iceland (Spacescape, 2014c).

2.2 A new bicycle strategy

The new bicycle strategy for Oslo 2015-2025 is based on the vision that Oslo is a bicycling city welcoming all bicyclists independent of age and skill. The vision is supported by three goals:

1) The bicycle share in Oslo should be at least 16 % in 2025.
2) The bicycle infrastructure in Oslo should be available, easy to manoeuvre and safe.
3) The Oslo inhabitants shall perceive Oslo as a safe and good bicycle city, suitable for children and the elderly.

Working towards these goals the strategy emphasises three areas of intervention:

1) The bicycle should be a part of life in the city and the physical urban environment.
2) There should be created a dense bicycle network with an “Oslo-standard”.
3) The whole of Oslo should take part in promoting bicycling.

As anticipated, infrastructure provisions are an integral part of the new bicycle strategy. In the next two sections I will describe the details of how the municipality are currently approaching the task they have assigned themselves.

2.2.1 A new bicycle network

In 1990 Oslo municipality and the Norwegian Public Roads Administration (NPRA) proposed a joint action plan for establishing a main network for bicycling in Oslo. The NPRA has responsibility for regional and national roadways, thus in the action plan they had responsibility for the parts of the bicycle network connecting to Oslo’s hinterland. As already described, this plan is the basis for the existing bicycle network in Oslo (Statens Vegvesen & Oslo Kommune, 2015). Following up on the directions in the Oslo bicycle strategy, Oslo municipality and the NPRA has joint forces once again to create a plan for the new bicycle network in Oslo. They released a consultation draft of this plan 31.08.15. The document takes point of departure in the evaluation of the existing network and the vision and goals formulated in the bicycle strategy.

The plan proposes two phases of development, one finishing in 2025 and the second has no time frame. The first phase consists of the provision of 80 kilometres of bicycle infrastructure. Together with the existing network this totals to 260 kilometres bicycle network, and results in 60 % of the population living within 200 metres of the network. The second phase brings the total up to 510 kilometres and the share of the population within 200 metres to 85 %.

The Spacescape evaluation of bicycle infrastructure had four main conclusions, these were that the bicycle network was to incoherent, that there were too frequent changes of designs, the facilitation in the inner city was too poor and the standard of designs did not hold a good enough quality. The concern about providing facilities in the Oslo city centre is emphasised in the new plan for bicycle network, and the plan states that the inner city is the priority in the first phase of the development (Statens Vegvesen & Oslo Kommune, 2015). This is reflected in the illustrations accompanying the plan, where it is obvious that the more suburban areas are downgraded. An extension of this is that the philosophy of the city centre as a focal point, recognized in the previous plan for the bicycle network is continued.

The map in figure 3 displays a draft of possible routes in the first phase of development. The drafted plan emphasises that none of the routes are decided and that it is only a principle drawing (Statens Vegvesen & Oslo Kommune, 2015).

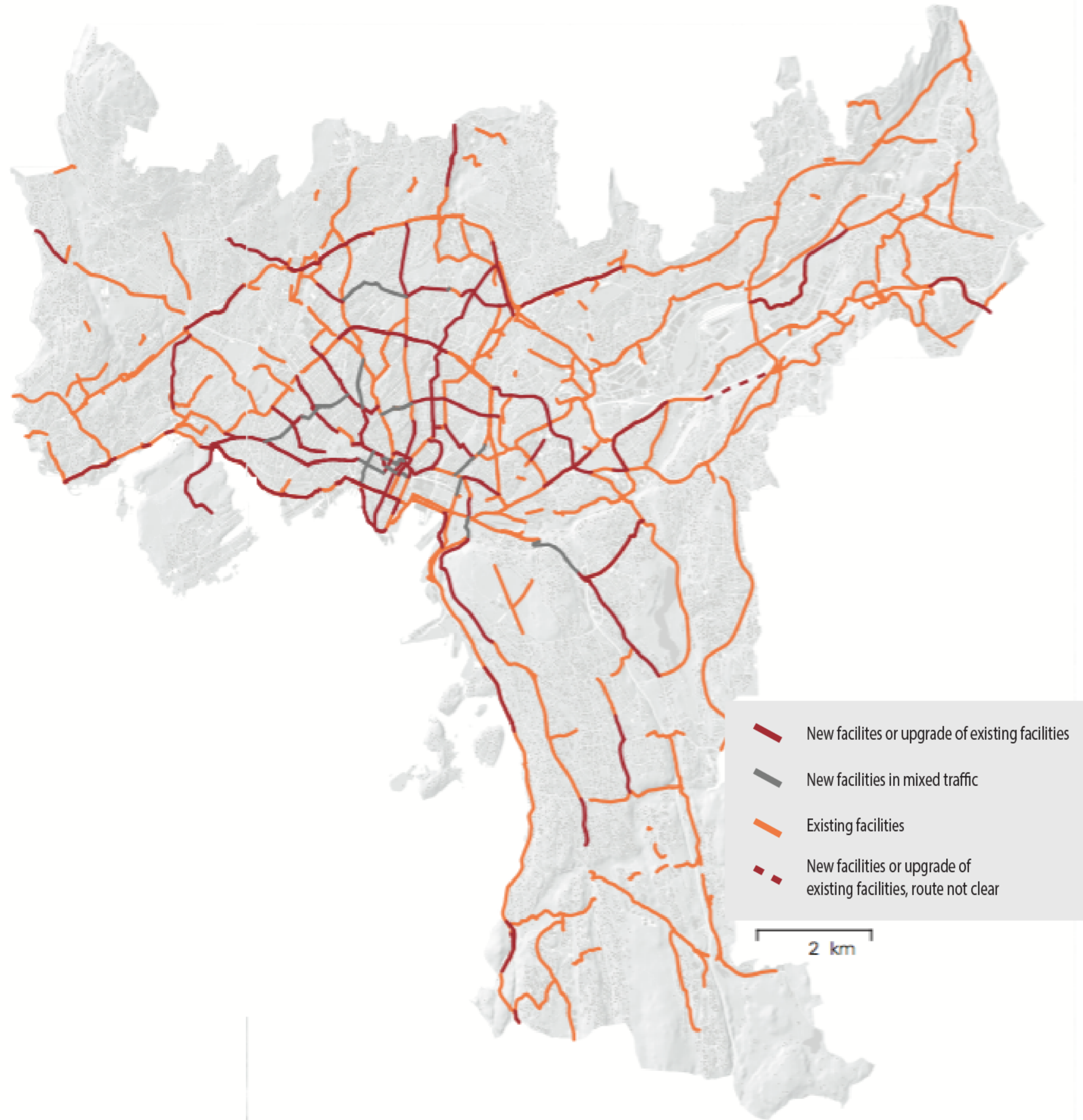


Figure 3 Draft first phase development of bicycle network (Statens Vegvesen & Oslo Kommune, 2015)

2.2.2 An Oslo-standard for bicycle infrastructure design

The Oslo bicycle strategy states that the Norwegian Public Roads Administration (NPRA) standards for bicycle infrastructure designs that have previously been used in Oslo are not sufficient in order to reach the goals set in the strategy. It suggests that the existing standards are mostly suitable for adult commuters, and is an important determinant for the over all dissatisfaction with the bicycle facilitation in Oslo. This has led to the conclusion that Oslo need it's own standard; a collection of bicycle infrastructure designs suitable for Oslo's visions and goals (Spacescape, 2014c).

The preliminary work with this development was started when focus groups from each the three bicyclist groups described above and a group of planners and “bicycle professionals” were asked to evaluate different features of infrastructure designs. The group of professional were thought to be able to distinguish between actual safety and perceived safety. The evaluation was done through a stated preference experiment, where the focus groups needed to agree on an evaluation ranging from very good to very bad for each design. The evaluation coincided with the suggestion that the current bicycle standard is not suitable for Oslo.

It concluded for one thing that drawn bicycle lanes with the current width requirements, are not an appropriate design for Oslo, but that bicycle lane with buffer zones or physical separation was perfectly fine. Further they the exercise found that separate bicycle paths were popular, both with and without division from pedestrians, although with separation was best. They found that the wider, the better, both for separate paths and bicycle lanes.

It also concluded that for the most parts there was little variation in the evaluation between the groups. However there was a marked difference in the perception of bicycling in environments with mixed traffic. Where the *Bikes no matter what*-group was more positive to both motorised traffic and pedestrians than the other groups and sometimes also the professionals. All groups did however agree that biking in a quiet street with low speed limits was a good solution.

The work with defining this standard has yet to be finalized, but the new plan for bicycle network described above make a suggestion of six possible designs. Of these four are dedicated to bicyclists; bicycle road with lanes for both directions, cycle track on each side of the road, bicycle street where other traffic is allowed on occasions and bicycle lane on each side of the road. In addition to these four designs it suggests shared pedestrian and bicycle paths, and bicycling in mixed traffic where the traffic load and speed is low. In addition to these six proposals they commit to creating a bicycle expressway.

Finally the bicycle strategy states that even though it will legally still be allowed to bicycle on the pavement in Oslo as in all other cities in Norway, bicycling on the pavement is not a good solution, an will not be a part of a the new Oslo standard (Spacescape, 2014c).

3 EXPLORING PREFERENCE FOR BICYCLE INFRASTRUCTURE

From the introduction we know that there are several determinants for bicycle use. It was established that the availability of proper bicycle infrastructure is an important determinant for whether or not people chose to bike, and that building bicycle infrastructure thus is an powerful measure for urban planners wishing to increase the bicycle share in their city. Further it was described that there exists no consensus on what type of bicycle infrastructure is the best choice, and that there is no gold standard to evaluate this because there are so many different concerns in the matter. However, one advance in this effort is to focus on what type of bicycle infrastructure the bicyclists themselves are most comfortable using; what infrastructure they prefer.

In this chapter I will explore how we can investigate preference for bicycle infrastructure and review what is already known about it. This will lead to the construction of a theoretical framework guiding this study.

The chapter will commence with a brief introduction of the mainstream approach to travel behaviour and how this approach can be used to get insights into preference for bicycle infrastructure among different types of cyclists. After that section I will review a selection of key literature employing this framework to investigate bicycle infrastructure. This will give insight to existing knowledge and serves as a backdrop for the findings of this study.

Following this I will explore behavioural theory as a way of expanding the presented approach in order to acquire more information on preference for bicycle infrastructure. The chapter is concluded with a summary where I introduce multi-dimensional bicyclist typologies to build a framework for this study based on the discussed theories and literature.

3.1 Preference and travel behaviour

Travel behaviour is the study of how people move through the physical environment. Traditionally the main objective of this effort has been to be able to forecast accurate patterns of travel (Handy, 2005). The patterns are depended upon estimating expected traffic flow when planning and engineering new infrastructure, or evaluating other changes to the system of transportation. The core of the forecasting exercise is to be able to say something about how traffic users makes different decisions and thus how people diffuse in space.

The most used theoretical approach in the field of travel behaviour is discrete choice theory (Handy, 2005). Opposed to earlier travel behaviour theories, which was often based on aggregate level data, the discrete choice theory takes travellers' individual choice into consideration. The inclusion enables the creation of more precise and flexible models (McFadden, 2002). The theory takes point of departure in the utility-maximizing framework known from economics and psychology, and is based on the assumption that all people make decisions that promote their self-interest e.g. that they choose the option that maximizes their utility.

The discrete choice theory states that the probability of an individual making a particular choice is dependent on the utility associated with that choice relative to all other available choices (McFadden, 2002). In relation to travel behaviour this implies that every traveller choose where and how to travel based on an evaluation of all costs associated with one alternative, compared to all other available alternatives. This process is referred to as route choice. Examples of costs are monetary costs (e.g. fuel, road pricing and public transport fare), time costs and convenience costs (e.g. traffic flow, parking facilities, shower facilities and perceived risk). The sources of these costs are termed attributes.

The utility of a specific alternative is described through a utility function, and it is the sum of these functions that make up the travel behaviour model. Each utility function has two terms. The first is an empirical term, which describes the relative importance of the attributes associated with the alternative. It is estimated from an aggregate of observed individual choices. The second is a "random" component. This accounts for all unexplained variations in taste and unobserved attributes (Handy, 2005).

The information on individual choices is collected from a sample of the population. There are two ways to obtain this information. One is to observe the individuals making a choice in their everyday life; per example counting how many people that choses one route over another. This type of collection is called revealed preference. It is the most conventional way of collecting information on choices (Wardman, 1988). The second method is to stage a scenario with two or more alternatives where the individual is asked what they would choose if they were faced with the choice in their everyday life. This is called stated preference. Stated preference is often easier to use, as it is a way to manufacture the perfect situation. It is easier to account for all attributes with the alternatives, and the researcher is at liberty to manipulate the alternatives so that they perfectly fit the research' objective. For instance the design frequently featured in the studies presented in the next section, trading time and convenience costs for a theoretically better facility e.g. one that is perceived to be safer. Such sophisticated situations are difficult to obtain in the “real world” with its many interrelations and whimsical nature (Pucher et al., 2010). The stated preference method is on the other hand criticised for not being able to convey accurate information, this might be because of bias or because the individual find the task difficult or overthink it (Wardman, 1988).

Revealed preference is a well-known expression in economic theory. It is not just a method for collection of data, it is a fundamental principle in choice theory (Sen, 1973). The principle was formulated by Paul Samuelson in 1938. It states that given two available choices, if one is chosen before the other it is also preferred above the other (Sen, 1973). This simple statement allows us to interpret the choices made by the individuals as expressions of preference. Combined with the rationale from the utility-maximizing framework, this means that choice is an expression of preference and is based on the individual’s desire to maximize utility. This enables conclusions in route choice modelling experiments such as: “Insofar females demonstrated preference, it was for off-road paths” (Garrard, Rose, & Lo, 2008, p. 58).

The random term in the utility function is an acknowledgement that not all people share the same taste e.g. that they perceive utility differently. To some degree this can be accounted for by dividing the sample into smaller groups that share similar characteristics. This act is called segmentation, and is often used in route choice theory. This initiative is why literature employing route-choice modelling is valuable in order to explore difference in preference between different types of bicyclists. The approach gives more nuanced information and makes it possible to compare preferences between different segments of the population. The most common form of segmentation within travel behaviour is using known categories such as gender, socio-economic status and frequency of travel (Handy, 2005; Julsrud, 2012).

To sum up, there is a mainstream approach to travel behaviour that relies on a utility maximizing framework. This means that information on individual choice between alternative routes is collected through revealed or stated preference, and is used to create utility functions which describe each individual’s willingness to pay for different attributes. The function should be interpreted as an expression of preference. Different segments of the population can have different perceptions of the utility associated with the same alternative.

As mentioned in the beginning of this section, this type of approach is traditionally used with the end goal of creating models of traffic flow. Today there are several studies published that employs this approach to investigate how travel environment affect choice, and thus preference for different features. Some of these studies investigate bicycle infrastructure, and in the next section I present findings from some of these.

3.2 Preference for bicycle infrastructure

In this section I will explore what the literature has to say about user preference for different types of bicycle infrastructure designs proposed for Oslo, and if there is evidence that different types of bicyclists evaluate the designs differently. Common for all of these studies is that they use route choice experiments to explore preference for different bicycle environments. Most of the studies draw on the utility-maximization framework described above, while some use descriptive statistics to illustrate differences between route attributes. All of the studies are concerned with non-recreational cycling e.g. cycling for transportation.

Through a stated preference survey, Stinson and Bhat inquired over 3000 individuals about bicycle route preferences for commuting purposes (Stinson & Bhat, 2003, 2005). They published two studies based on this survey, where the latter added value by categorising the respondents based on their level of experience. The segmentation was based on their self-reported level of experience. Preliminary descriptions of the segments showed that in general females reported to be less experienced with bicycle commuting than men (Stinson & Bhat, 2005).

A result reported in both studies, was that the respondents over all preferred routes with bicycle facilities such as bike lanes and paths, to those without any facilities, and that they preferred bike lanes to separate paths. It is not stated if the path was intended shared with pedestrians (Stinson & Bhat, 2003, 2005). The second study found that experienced cyclists are less concerned with the level of separation from motorised transport than inexperienced cyclists. For experienced cyclists time savings was the most important factor (Stinson & Bhat, 2005).

In 2007, Hunt and Abraham published a study based on a stated preference survey among 1128 cyclists. The survey was set up to investigate how respondents would trade off different type of bicycle infrastructure for time costs and the convenience of showers and secure parking at destination. They used socio-economic factors, level of experience and comfort cycling to categorize the respondents (Hunt & Abraham, 2007). Over all they found that time spent cycling in mixed traffic without any bicycle facilitation is more onerous than cycling on a bike path shared with pedestrians, and that both is more onerous than cycling on a bike lane.

They also found that time spent biking in mixed traffic becomes less onerous as the level of comfort and experience increases (Hunt & Abraham, 2007). Both of these findings resonate with the findings already described from the Stinson and Bhat studies (2003, 2005).

A contemporary study also concludes that bike lane is the most preferred bicycle facility (Tilahun et al., 2007). Equal to the three previously described studies, it employs a stated preference survey to explore preference for different cycling environments, and it can be found that the respondents have higher willingness to pay for bike lanes than other facilitation. The runner up was no street parking followed by off-road paths.

However this study did not find any difference in preference between cyclists and non-cyclists, which means that there is no difference between the respondents with much experience cycling and those inexperienced: "The cyclist variable, which indicates if the subject uses bicycling as their main mode at least during summer, is highly insignificant; indicating that preferences are not dictated by experience at least in this SP context." (Tilahun et al., 2007, p. 298). They did, on the other hand, find a difference in preference between the genders, in that the female respondents had a higher tendency to choose the facilities that are perceived safer than men (Tilahun et al., 2007).

In a recent case study from Dublin, almost 2000 people working in the city centre responded to a stated preference survey inquiring about how different type of bicycle infrastructure would affect their choice to bike to work (Caulfield, Brick, & McCarthy, 2012). When asked directly about this, 74.1% said that more off-road cycle tracks and greenways would encourage them to cycle to work, and 56.4 % said that more bicycle lanes would do the same.

The results from the stated preference experiments varying travel time, exposure to traffic and other costs, confirmed the respondents' initial suggestions, showing that the respondents over all were more likely to choose routes with an off-road cycle track, followed by a greenway (Caulfield et al., 2012). Cycle track was not featured as an option in the previous studies, but the fact that greenway was preferred to bicycle lane diverge from the previously described findings.

The respondents were least likely to choose to bike in mixed traffic and in a shared bus and bike lane. The respondents were divided into segments based on their reported confidence level bicycling in order to explore if the probability varied between segments, but the study found no direct correlation between confidence level and preference for infrastructure (Caulfield et al., 2012).

In Portland, Oregon 248 bicyclists were stopped and asked to answer a survey concerning two new types of buffered bicycle facilities. One is a cycle track, where the facility is buffered from traffic with a shy zone and parked cars, the second is a solution where one of the lanes in a two-lane roadway is designated for bicycles, separated from traffic in the opposite direction by a drawn shy zone. The purpose of the new facilities is to make the bicyclists feel more comfortable and to increase biking (Monsere, McNeil, & Dill, 2012).

The survey both inquired about the bicyclists' experiences with the new facilities and presented them with a route choice dilemma. Based on the bicyclists' reported perception, the study concluded that women were significantly more likely to agree to that the cycle track made it safer to bike. There was no difference between the genders on the buffered bike lane.

When separating the respondents into those who feel comfortable cycling in a bike lane with heavy traffic from those who feel uncomfortable, there was a significant difference in how many that felt they were safer with the new facilities. 69% of those who felt comfortable compared with 84 % with those who felt uncomfortable, indicating that the effect of "segregated infrastructure" is larger with the group who feel uncomfortable bicycling. When asked to trade travel time for better facilities, 59 % of the respondents would trade to bike on the cycle track and 71 % to bike on the buffered lane (Monsere et al., 2012). The findings suggest that the buffered lane is the more attractive option.

While the previous study sported some sort of hybrid design, the other studies presented above were based on stated preference experiments. In the following the findings from three relevant revealed preference studies are presented.

In 2008, a study was published where the main objective was to investigate whether female bicyclists are more likely than male to use bicycle facilities with greater separation from motor vehicle traffic. The proposed mechanism at work is that women are more risk averse than men, and would thus prefer more segregation. No other indicators than gender were used to categorize the bicyclists in this study (Garrard et al., 2008).

The study utilised a revealed preference design where the data was collected through observations. The initial observations indicated that females showed preference for on-road bicycle lanes and no facilities, compared to off-road facilities e.g. bicycle paths. However, when the researchers corrected for distance to destination, female bicyclists showed significant preference for off-road facilities compared to both on-road and no facilities. It was also reported that female bicyclists showed no preference for on-road lanes over roads without any facilities (Garrard et al., 2008).

The two next studies are revealed preference experiments. They are both based on the same GPS data, collected from 166 bicyclists in Portland in 2007. The GPS recorded all bicycle trips made by the bicyclists during a period of seven days. The respondents complemented the data with trip purpose, weather information and an evaluation of what attributes was important for them when choosing this route. The participants were experienced cyclists, as most of the participants biked five days a week or more, and the authors infer that they are likely to be confident cyclists (Broach, Dill, & Gliebe, 2012; Dill, 2009).

The study from 2009 investigated how bicycle infrastructure can increase levels of bicycling. By showing that half of the miles travelled was on some sort of bicycle infrastructure, while only 8 % of the road network in Portland have such facilitation, it was concluded that people go out of their way to use bicycle infrastructure. Yet the study also showed that when the bicyclists were asked to rate the importance of factors when choosing bicycle route the most important factor was to minimize total distance travelled. Avoiding streets with lots of traffic was rated second, while riding on a bike lane was the third most important concern, and riding on an off-street path the sixth, followed only by avoiding hills (Dill, 2009).

The most recent study uses the GPS data to develop a route choice model. The model showed that the bicyclists showed highest preference for separated paths (Broach et al., 2012). This differs from when the bicyclists themselves rated the importance of factors in the previous study. It also showed that bicycle lanes were only preferred, when there was no option of a low-traffic neighbourhood street (Broach et al., 2012). This ranking corresponds to the ratings presented in the previous study. The fact that the two studies give different indications highlights the importance of employing different methods. The study concludes that even confident and experienced bicyclists prefer routes that reduce exposure to motorised traffic (Broach et al., 2012).

Following I will give a summary of the findings described in this section, and compare them to those found in a comprehensive international review on different interventions to increase bicycling published in 2010. The review evaluated several types of interventions, one of which is bicycle infrastructure. Several revealed and stated preferences studies were included to investigate how bicycle infrastructure was used and perceived, and the studies presented above published before 2010 were included in the review (Pucher et al., 2010).

The nine studies featured in this section offers no clear indication as to what infrastructure is the most desirable. Several studies conclude that bicyclists have higher willingness to pay for bicycle lanes than bicycle paths (Hunt & Abraham, 2007; Stinson & Bhat, 2003, 2005; Tilahun et al., 2007), while others conclude that separate paths is more preferable (Broach et al., 2012; Caulfield et al., 2012). However, all conclude that some kind of facilitation is preferred to biking in mixed traffic, and this is consistent with the review. This being said, it is evident that there are measures to be made in order to make biking in mixed traffic more attractive, such as signage, removing street parking and creating bicycle boulevards (Broach et al., 2012; Tilahun et al., 2007).

Cycle tracks was only featured in one of the studies, where it was rated the most preferable type of facility (Caulfield et al., 2012). The review included several studies investigating this type of infrastructure. They found two additional studies rated that cycle track as more desirable than striped lanes, and studies concluding that building cycle tracks had increased the number of bikes on the roadway in both London and Copenhagen, respectively 58 % over 3.5 years, and 20 % in an before-after study (Pucher et al., 2010).

There is also conflicting evidence as to whether or not there is differences in preference between experienced and inexperienced bicyclists. When confidence, level of comfort and experience is likened, three studies found evidence that there was a difference (Hunt & Abraham, 2007; Stinson & Bhat, 2003, 2005), while two found no evidence to support the claim (Caulfield et al., 2012; Tilahun et al., 2007). The tendency described in the studies showing difference in preference is that less experienced bicyclists prefer more segregated solutions. This tendency is also described in Monsere et al., without being properly backed with significant evidence (2012). Broach et al. suggest the opposite tendency in their study. They are not able to give evidence of any difference between experienced and inexperienced bicyclists, as their study almost only includes experienced bicyclists, but they find that they prefer segregated infrastructure and as little exposure to motorised traffic as possible (Broach et al., 2012).

The review stated that there was evidence to suggest that experienced cyclists prefer bike lanes to off-road paths (Pucher et al., 2010). This based on six studies, including the studies previous presented by Stinson and Bhat (2003), Hunt and Abraham (2007) and Tilahun et al.

(2007). However, as described in a previous section, I do not recognize the finding in Tilahun et al., and two of the studies published after the review was completed nuance this claim (Broach et al., 2012; Caulfield et al., 2012).

Further, several findings indicate that female and male bicyclists have different preferences, and that females prefer infrastructure with less exposure to motorised traffic (Garrard et al., 2008; Monsere et al., 2012). This is also emphasised in the review referring to two additional studies (Pucher et al., 2010). One frequently proposed explanation for this tendency is that women are more risk averse than men (Garrard et al., 2008).

However another explanation for this might be that women in general are less experienced bicyclists than men, as suggested in Stinson and Bhat, and that the difference in preference might just be an interaction effect. Thus we do not know whether there is a difference because they are inexperienced, or because they are *women*. As far as I can see this is not accounted for in the studies. In Garrard et al. the study was based on study was based on observations, and there was no possibility to investigate the level of experience among the bicyclists (Garrard et al., 2008).

All studies presented above, except Garrard et al., recorded the age of the respondents. Some of the studies consider age as a factor affecting preference, but no study report any significant difference, there is no mention in the review about age affecting preference (Pucher et al., 2010). Something that is worth to make a note of is that all of the studies have adult respondents.

This section makes no mention of shared space solutions, and neither is there any direct indication of the effect of bicyclist sharing space with pedestrians on pavements or on bicycle paths. Pucher et al. concludes in their review that they were unable to find any published studies on shared space solutions (in the review referred to as complete streets), neither do they discuss or describe interactions with pedestrians as a factor affecting preference (Pucher et al., 2010).

3.3 Expanding with behavioural theory

As can be read in the previous review section, Garrard et al. propose that the mechanism at play for women preferring off-road paths and designs separating bicycles and motor vehicles is that they put higher value on the feeling of safety (Garrard et al., 2008). The utility framework most of these studies are based on, can neither confirm nor deny the proposal. This example illustrates a well-known critique of this approach to travel behaviour (Handy, 2005). It gives evidence that there exists a difference in preference between different segments of bicyclists e.g. a difference in utility, but is unable to say anything about the reason why. Exploring such causal relations is valuable for planning purposes. It allows for more nuanced and precise interventions, and makes it easier to predict the effect of the interventions planned.

I this respect it seems important to find a theoretical approach that can help explain the different choices made. A candidate is behavioural theory and the theory of planned behaviour. Handy emphasis in her assessment on literature exploring the relationship between land use, transportation and physical activity that behavioural theories can be a valuable supplement to the utility-maximizing framework when exploring the relationship as “they are more explicit about the specific variables that explain behaviour, and less about the mechanisms by which these variables act on behaviour” (2005, p. 13). Using the conceptual framework from travel behaviour I interpret the connection between the two approaches to be that while the utility-maximizing framework is concerned with how route choices are based on each person maximizing their utility, the behavioural theories are concerned with what factors the individual believe will maximize their utility.

As anticipated the theory of planned behaviour is an important theory in the field of behavioural theory, and it is one of the most cited and applied theories. It is developed by Ajzen and is an extension of the theory of reasoned behaviour by Ajzen and Fishbein. The theory postulates that “behaviour is a function of salient information, or beliefs, relevant to the behaviour” (Ajzen, 1991, p. 189). This means that beliefs about behaviour are a determining factor for actual behaviour.

Ajzen distinguishes between three types of beliefs; normative beliefs, behavioural beliefs, and control beliefs (1991). Normative beliefs are concerned with whether or not people important to you would approve of a type of behaviour, and furthermore if you yourself approve of the behaviour. Control beliefs are beliefs about how likely it is that you are able to go through with the behaviour. For example if you believe you are able cross an intersection running a red light, you might just try it. These beliefs are based on passed experiences and second-hand beliefs. The last type of beliefs is behavioural beliefs. These concerns are about what outcome you are likely to achieve with the behaviour. Seen in connection with bicycling, examples are if you believe you bike faster by choosing to bicycle in one environment, if you believe you could get killed or maimed bicycling in another environment, or if you believe you are safer using exercise equipment.

3.4 A richer approach to preference

The segmentation in the articles reviewed above is mainly one-dimensional. This means that the researchers have distinguished between different bicyclists using one characteristic. There are three dominating characteristic, gender, experience bicycling and how comfortable you are bicycling. No clear conclusions can be derived from the travel literature as to whether or not these characteristics determinant preference for different infrastructure designs, but that the tendency in the studies that were able to report such a finding was that women and those less experienced prefer segregated solutions. Only using one dimension to describe bicyclists reduces the possibility to capitalize on the diversity in the bicyclist population, and the reported findings are not necessarily robust enough to use when planning bicycle infrastructure. There is more room for interpretation and hypothesising about different mechanisms at play if the segments are richer.

This challenge is addressed in the study *What is your type: A multidimensional cyclist typology* by Damant-Sirois et al. (2014). They propose using cluster analysis as a methodological tool to explore diversity in the bicyclist population. The use of this method to investigate the bicyclist population is also suggested Handy et al. in their assessment of research needs and challenges to promote bicycling (2014). They emphasise the possibility for using this method as a way of employing multiple behavioural indicators to explore mobility patterns.

Using cluster analysis to create a multidimensional typology is a way of creating richer segments of bicyclist. Taking inspiration in behavioural theory when choosing what dimensions that should be included in such an analysis will hopefully provide more robust results when investigating preference for bicycle infrastructure designs among different bicyclists.

4 METHODOLOGY

The aim of this study is to investigate different bicyclist types in Oslo and their use of infrastructure. This calls for an exploration of patterns within a population, and for this purpose a quantitative approach is valuable. I will base the analysis on existing survey data, as this gives me access to a sample size and is more efficient given the short duration of this project.

4.1 Research design

4.1.1 The Safety in Numbers-project

The data material featured in this study is collected through surveys in connection with the research project “Safety in Numbers” (SiN). This project is a collaboration between Norwegian Institute of Transport Economics (TØI) and researchers from Lund University, Aalborg University, the University of Bath and SWOV Institute for road safety research in the Netherlands. The project is designed to investigate potential mechanisms of the safety in number effect (Forskningsrådet, 2014). Where, the proposed effect is that more bicyclists in a traffic environment provide a safer environment for the individual bicyclist (Jacobsen, 2003; Robinson, 2005).

The surveys featured in the project, were conducted in 2013 and provide longitudinal input to investigate changes in one city throughout the bicycle season, and cross-sectional input to investigate the effect of different infrastructure designs and traffic culture. The survey was carried out at three different points in time in Oslo (April 2013: T1, June 2013: T2 and September 2013: T3), and at one point in time (October 2013: T3) in Aalborg and Gothenburg (Fyhri & Sundfør, 2013).

Bicyclists were recruited on the street in different locations in the city, and asked if they could complete a short questionnaire, approximately 4 minutes. Those who agreed to participate were asked at the end of the questionnaire if they would like to participate in a follow-up survey at home the same day. After completing the home questionnaire participants recruited at T1 and T2 were asked if they would participate in a new survey in six weeks. The respondents were informed that each time they participated (T1, T2, T3), they would qualify to be part of a lottery for a gift certificate of 5000 NOK.

The three versions of the home questionnaire are almost identical, but some of the control questions vary. The first one includes questions about the respondent’s personality, whereas the second and third do not include these questions but are otherwise identical. The third includes questionnaire includes question regarding political affiliation. The given priority of the versions is first, third then second. Thus the ones recruited at T3 were given the first questionnaire, the ones recruited at T2 were given the first and the third and the ones recruited at T1 were given all home questionnaires. All respondents answered the field questionnaire. The survey design is illustrated in the diagram below.

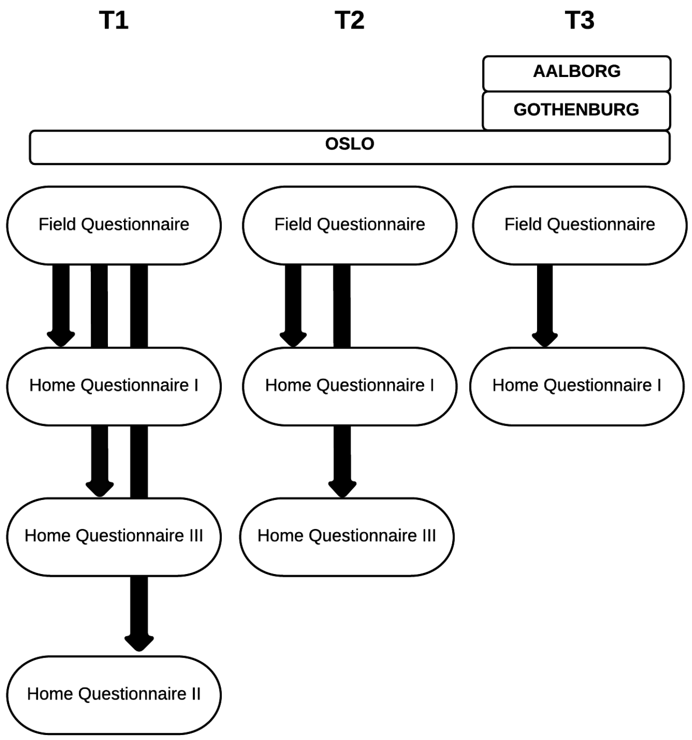


Figure 4 Survey design in the Safety in numbers project.

The questionnaires were designed to investigate the safety in numbers effect (Fyhri & Sundfør, 2013), hence the predominant concern was interplay between different traffic users and perceived safety in different situations. In addition to inquiries about interplay, the bicyclists were asked what type of bicycle infrastructure they had biked on during the past week, how frequently they bike and asked about their use of bicycle equipment and behaviour. The field questionnaire asked specifically about the bicyclists' experience and behaviour that day, while the home questionnaire asked the respondents to review their experiences from last week. The field questionnaire and the first home questionnaire are included as appendix A and B.

As the respondents are asked to describe their behaviour for a past period of time, the design is revealed preference. They reveal what they chose to do, not what they would have chosen.

Yet, it is self-reported and not as robust as some other revealed preference design, for instance where GPS or observations are used (Broach et al., 2012; Dill, 2009; Garrard et al., 2008). When asked to describe their behaviour the past week, or even the same day, the respondents might remember wrong or even answer what they believe they would choose rather than what they actually chose, thus to some degree actually answering their stated preference.

4.1.2 Adapting the research design

As described in the previous section The Safety in Number project includes many elements. Not all of these were relevant for the present study. As Oslo is the empirical context for this study, it made sense to only include data from Oslo. The objective of the present study was to describe a population. Hence a cross-sectional design was desirable. The primary areas of interest in the present study, use of infrastructure and personal traits, are characteristics that are not believed to change with the season. This means that the longitudinal dimension of the original design was not especially relevant for the present study. In order to get the largest possible sample size, and provided that the time of the season is not of importance, the sample featured in this study is all Oslo respondents who answered the *first* home questionnaire.

Employing survey data from the Safety in Numbers project allowed access to a large sample of bicyclists, which could not have been obtained over the short period of time available for this study. Thus, using the survey data opened doors, but only ajar. The consequence of using the surveys as data material was that the scope and quality of the study was dependent on the questions asked in the questionnaires and how the surveys were designed. In the following I will discuss how the current study was influenced by the original research design from the Safety in Numbers-project.

The first research question, “What are the different types of bicyclists in Oslo?” calls for an exploration of the individual character traits among the respondents, ideally to create a multidimensional bicyclist typology. In the original research design, individual characteristics of the bicyclists were needed for two purposes. First as control variables moderating the safety in numbers effect between different points in time, for example could one assume that more inexperienced cyclists during summertime obscure the effect. Second it was used to delineate different traffic cultures between the three cities (Fyhri & Sundfør, 2013).

In the present study individual characteristics were needed in order to create a bicyclist typology. The typology should be able to inform planning decisions, and preferably display difference preference for different type of bicycle infrastructure. The appropriate characteristics were delineated in chapter 3, and were decided by theoretical considerations and earlier comparable studies. They should preferably include demographics, bicycling experience and perceptions and beliefs about bicycling.

The survey included questions about gender and age. Unfortunately, neither of the questionnaires included any information on socio-economic status, employment or education. These are all indicators commonly used as control variables in similar studies (Damant-Sirois et al., 2014; Stinson & Bhat, 2003, 2005). They could be interesting to include in order to study diversity in the bicyclist population, and to manage the objectives in the Oslo bicycle strategy more directly, for example say something specific about the preferences among students.

Several questions in the questionnaire described how the bicyclist bikes. They were asked about intensity and the use of equipment. These types of questions gave insights to what bicycling entails for the individual bicyclists, and can be interpret as their perception and belief about bicycling. The use of safety equipment can indicate that they believe bicycling is dangerous or that they feel uncomfortable in the current environment and the use of exercise equipment and bicycling intensity can indicate that they think of bicycling as exercise. The questionnaires included a question about the bicyclists' perception of safety when biking. They also included a normative statement concerning the bicyclists' right to break traffic regulations. Both of these questions are directly related to the bicycle experience and the perception of biking.

The second question could prepare for a route-choice experiment, where the typology was used to create segments of bicyclists, comparable to some of the studies described in section 3.2. However, the research design in the Safety in Numbers-project is not so much concerned with route-choice dilemmas. In the original research design, exposure to different types of bicycle infrastructure is proposed as an explanatory factor in the safety in numbers effect. Thus, the surveys were designed to give information about infrastructure exposure on an aggregate level, and not the bicyclists' choice between different infrastructure types.

One question in the survey inquired specifically about use of various types of infrastructure. It asked bicyclists to divide their time bicycling between different types of infrastructure, respectively the pavement, separate bicycle paths, bicycle lanes and mixed traffic. It gives no information about what alternatives were available for the bicyclist. So while this question can't say anything about specific choices between facilities, each bicyclist' infrastructure-distribution can give indication for preference. The bicyclists can choose between routes with different facilitation, and can choose to bike on the pavement instead of in mixed traffic. This being said exposure in the immediate environment is probably an important explanatory factor.

This means that there was no possibility to conduct a traditional route-choice experiment where trade-offs between different types of facilities and other factors can be studied. Instead the frequency of use of a specific type of infrastructure was used as a dependent variable to see if it was affected by belonging in the typology.

In addition to the reports on frequency of use the bicyclists were asked about how unsafe they felt bicycling in different infrastructure. This included separate bicycle paths, bike lanes and in mixed traffic. When this information was used together with information on the respondents' general level of insecurity it gave an indication on the value of different facilities for different users. This construct was also included as a dependent variable.

Together the two indicators, use of infrastructure and valuation of infrastructure, could say something about the bicyclists' preference different types of infrastructure designs. What designs are however limited to those described in the survey; separate bicycle path, bike lanes, pavement and mixed traffic.

All indicators and variables are more thoroughly discussed in chapter 5.

4.2 Analytical design and statistical methods

In this section I provide an overview of the analytical design before presenting the relevant statistical methods and the considerations made before applying them in the study.

4.2.1 Analytical design

The first step of the analysis was to select indicators suitable as dependent and explanatory variables. This process was initiated by the theoretical discussion in chapter 3 and decided through the adaptation of the research design, and is described in the previous section, 4.2.2. The next step was to create operational variables. This process is described in detail in chapter 5.

Five of the explanatory variables were included in a K-means cluster analysis. The variables included in this analysis were: Use of safety equipment, use of exercise equipment, biking intensity, perceived insecurity and liability to break traffic regulations. The result of this analysis was a five-dimensional bicyclist typology. The remaining three variables, gender, age and experience, were used to qualify the cluster solutions and offer additional information about the bicyclist types.

After concluding the cluster analysis the resulting typology was tested against each of the dependent factors, frequency of use and valuation of different infrastructure designs. This was done using ANOVA with Tukey HSD post hoc test. The effect size (eta squared) was calculated to determine how much of the variation in the dependent variables that could be explained by the bicyclist typology.

The cluster analysis was conducted using the programming language and software R, as it provided a possibility for assessing the validity of the clusters that was not available (as far as I know) in SPSS. The clustering solutions created in R were later exported to SPSS in order to proceed with ANOVAs.

4.2.2 Statistics and hypothesis testing

Statistical method can be divided into two branches, descriptive statistics and inferential statistics. Descriptive statistics is used to explore the empirical patterns in a sample such as means and shares, while inferential statistics enables the researcher to generalise from the sample to a population (Bakeman & Robinson, 2005).

Inferential statistics use statistical decision theory or hypothesis testing to assess whether the pattern within a sample can be ascribed to chance, or if it accurately reflect the pattern in a whole population (Bakeman & Robinson, 2005).

Different inferential methods use different statistical tests to decide whether the relations are significant, and thus can justify generalisation from sample to population. The relationship is often described through dependent and independent variables, where the independent variables are tested to see if they account for variations in a dependent variable (Bakeman & Robinson, 2005). In this study I employ the F-test, which is inherent in an ANOVA and the Chi-square test. Both of these tests will be briefly explained in the two following sections. The point of departure for these tests is nevertheless the same, the null hypothesis. This hypothesis is contrasted by the alternative hypothesis and these are constructed so that they are mutually exclusive and exhaustive (Bakeman & Robinson, 2005).

The standard error is a function of the sample size and standard deviation, and is used to calculate the probability for accurately rejecting the null hypothesis. This probability is called the p-value. In order for the result to be statistically significant the p-value need to be less than the chosen alpha-level. In social sciences it is common to operate with an alpha level of 0.05 e.g. that the probability of accurately rejecting the hypothesis is 95 % (Bakeman & Robinson, 2005). This level is employed in all statistical tests in the present study.

4.2.3 ANOVA

The name ANOVA is abbreviation for ANalyse Of VAariance. It is appropriately used when you have one categorical independent variable with at least two categories and wish to test it against a quantitative dependent variable. The method compares the means of two or more groups. It does so using the sample's F-distribution and a F-test, testing how the variance varies between the samples. The test yields if there is significant difference between the means of two or more categories (Bakeman & Robinson, 2005).

The ANOVA is an omnibus test. This means that it tests if there exists a difference between any of the categories. If at least two of the categories are significantly different, the test will be significant. In order to investigate which of the categories actually are significantly different one applies a Tukey Honest Significant Difference test (Bakeman & Robinson, 2005).

The ANOVA can also say something about how able the independent variable is to explain the variation in the dependent variable e.g. how much explanatory power the independent variable has in describing the phenomena described by the dependent variable. This is called effect size, and is measured by eta squared. The size of the effect can be interpret through Cohens' rule of thumb, where eta squared = .01 is considered a small effect, =.06 a medium and =.14 a large effect (Cohen, 1977).

In this study ANOVA is used to test if the difference in use of infrastructure and experience of safety on different infrastructures is affected by bicyclist type belonging.

4.2.4 Cluster analysis

A cluster analysis is an exploratory tool for organising data into groups when the groups are not known in advance. The analysis provides no statistical tests in itself, but is a method of organising information for future use. Breaking large amounts of data down in more manageable segments is essential for interpreting and understanding complex phenomena and this makes the tool useful for several different disciplines within both natural and social sciences (Landau & Ster, 2010).

4.2.4.1 Selecting the right clustering algorithm – K-means partitioning

The clusters are created based on a pattern of similarities and dissimilarities in selected independent variables. There are different numerical methods of exposing a pattern, thus several different approaches to doing a cluster analysis. The most common methods are hierarchical methods and partitioning methods. Common for all approaches is that they aim to maximize the differences between the groups while minimizing the difference within the groups under given constraints, such as the number of clusters and iterations (Mooi & Sarstedt, 2011).

In this study the K-means method was employed. The method is recommended when there are more than 500 objects and none of the variables are categorical (Mooi & Sarstedt, 2011). This is the case in this study. It was also the method employed by Damant-Sirois et al. to create a somewhat similar bicyclist typology (Damant-Sirois et al., 2014).

The K-means method is a partitioning method that is based on reducing within-cluster variation. As the classification suggests it uses an algorithm that starts with splitting, or partitioning the objects into a predetermined number of clusters, k. This is done by creating random cluster centres and assigning each object to the centre closest, measured in Euclidean distance. Euclidean distance is a metric measure between two points in space (A and B) and is given by $d_{Euclidean}(A,B) = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$. In the next step the geometric cluster centre for each cluster is computed and the initial random centre is replaced. The objects are then reassigned to the cluster centre that is closest, and a new cluster centre is computed. This process continues for a predetermined number of iterations or when convergence is achieved (Mooi & Sarstedt, 2011). The algorithm returns each object's group belonging and the location of the cluster centres in a multidimensional space, where the number of dimensions is equal to the number of independent variables included in the study. The location of the cluster centre is given as coordinates, where each coordinate represents the mean value of all observations within the cluster for the given dimension.

4.2.4.2 Presumptions for using K-means partitioning

A requisite for the variables included in the analysis is that that they do not have to high collinearity as this will interfere with the clusters' potential capacity to explain a dependent variable. The threshold is a Pearson's r of 0,9 (Mooi & Sarstedt, 2011). The collinearity was tested in a bivariate correlation analysis.

The use of Euclidean distance in the K-means algorithm presumes that the variables are ordinal or continuous and that the absolute variance of each variable is equal, thus the variables should have the same range. This is not the case for variables intended to be included in the present analysis, but it is resolved by standardizing the variables before they are used in the analysis.

Another requisite for using the K-means method in a cluster analysis is that the number of clusters needs to be decided by the researcher before running the algorithm. Making a decision about what number of clusters is relevant for the analysis is not always straightforward and is an essential part of the analysis. The evaluation is important, as it will have great influence on the results.

4.2.4.3 Determining the relevant number of clusters – a discussion of validity

An important determinant in the effort to decide the relevant number of clusters is cluster validity. Cluster validity refers to the quantitative evaluation of how well the clusters exposes the "hidden" structure in a data set (Halkidi, Batistakis, & Vazirgiannis, 2001). There are three different approaches to determining validity, namely evaluating external criteria, internal or relative criteria. External criteria implies evaluating the cluster solution based on an external pre-specified structure, per example a probability distribution, while the internal criteria approach only evaluates how inherent features in the data set fit the cluster solution (Charrad, Ghazzali, Boiteau, & Niknafs, 2014). When evaluating relative criteria the basic idea is to compare a selection of indices between cluster solutions. The solutions are generated using the same algorithm but varying the parameter values. As already established, the "number of clusters" is an input parameter in the K-means algorithm. This means that the approach gives information on what number of clusters that optimizes cluster validity, e.g. the number that most accurately exposes the underlying structure in the data set. The two other approaches imply no such comparison, and therefore does the relative criteria approach to cluster validity seem to be the most efficient when investigating the relevant number of clusters (Charrad et al., 2014).

There is a wide range of indices that has been proposed for evaluating relative criteria (Charrad et al., 2014). Most clustering validity indices take point of departure in measures on separation between clusters and compactness within clusters. But they also evaluate the geometric and statistical properties of the data set, the number of objects and measures of similarity and dissimilarity. Charrad et al. have gathered 30 such indices in a R-package, *NbClust*, designed especially for determining the number of relevant clusters, given a specific clustering algorithm, such as K-means (2014). The indices are collected from previous R-packages and a completive literature review and create an exhaustive list of validity indices relevant for evaluating the optimal number of clusters.

NbClust is designed with an option to test for all indices at once, generating a suggestion on the optimal number of clusters (Charrad et al., 2014). More often than not, different indices will suggest different number of clusters as optimal. In that case NbClust proposes to determine the optimal number through majority rule, meaning that the number of clusters most frequently suggested is the optimal number (Charrad et al., 2014). The integral use of the majority rule implies that NbClust can be used to determine the optimal number of clusters without having sufficient knowledge about the capacity of the different indices to be able to weight their suggestions.

The indices available, their authors, the year they were first proposed, as well as a short description of their features are presented in appendix C.

Summing up, a quantitative evaluation of cluster validity based on relative criteria is an important determinant when choosing the relevant number of clusters in the analysis because it finds the number of clusters that is optimal in order to best expose the underlying structure in a data set.

While the quantitative evaluation of cluster validity is an important feature, it should not be the only determinant when deciding on the number of clusters. As mentioned in the very beginning of this section, cluster analysis is an exploratory tool used to make data more manageable. This gives that an important measure of the quality of the analysis is whether it is possible to use the results for a practical application. This objective is reflected in literature written about cluster analysis for academic disciplines with this orientation (Mooi & Sarstedt, 2011).

In the chapter on cluster analysis in Mooi and Sarstedt's *A concise guide to market research* the authors never even mention quantitative measures of validity, but emphasise on qualitative measures, such as criterion validity and face validity. Criterion validity is also called construct validity, and refers to whether or not the clusters "make theoretical sense" in relation to concepts outside of the analysis (Bryman, 2012; Mooi & Sarstedt, 2011). Face validity is basically asking experts or even just ordinary people if the result makes sense to them (Bryman, 2012). In the chapter Mooi and Sarstedt point to the body of literature on segmentation for marketing purposes, and proposes coherence with criteria developed in this respect as a measure for the validity of a cluster solution. Examples are that:

- The segments should be large enough to serve.
- The segments can be distinguished conceptually.
- The segments are parsimonious, meaning that only a small set of substantial clusters should be identified in order for them to be managerial meaningful.
- The segments should be relevant for the actor's objectives (2011).

This account illustrates that there is different approaches to cluster analysis as a method, and thus how to determine the number of clusters. Literature from disciplines such as data mining and statistics emphasises the numerical aspect of cluster analysis, and asserts that the only measure of validity is how well the results reflect the actual structure of the data set (Charrad et al., 2014; Halkidi et al., 2001; Landau & Ster, 2010). While other disciplines focus on the practical aspect of the method e.g. that cluster analysis is a tool enabling future applications, and emphasises measures of validity connected to how well the result is suited for such objectives (Mooi & Sarstedt, 2011).

Returning to the question at hand, how to best decide the number of clusters, it becomes evident that there sometimes is a difference between the relevant and the optimal number. Thus, when deciding the relevant number, one is prompted to balance quantitative measures of validity against the objective of the study. In a sense it becomes a compromise between conveying the structure of the data set as accurately as possible and creating segments that are conceptually and theoretically recognizable, and thus allows for further analysis.

As a concluding remark in the discussion on the relevant number of cluster, is it appropriate to revisit the capacity of the K-means clustering algorithm and it's defining concern of minimizing within cluster variation. Typically, within cluster variation is expected to decrease as the numbers of clusters increases e.g. one cluster per observation equals zero within group variation. This means that a larger number of groups *will* contain more diverse information. This fact is independent of the evaluation of quantitative evaluation of validity, as the indices used for this concern the *optimal* point. Those indices occupied with within cluster variation will typically indicate the point where change in variation is at it's highest. In other words, where each additional group gives less additional information than earlier (Landau & Ster, 2010). This will for instance mean that even though the quantitative evaluation of validity suggests that two clusters is the optimal number of clusters for the data set, three or more clusters will contain more information and could thus be deemed more relevant for the study.

4.2.4.4 Carrying out the cluster analysis

The first step of the analysis was to import the data set with the selected variables to R. Following the NbClust package was employed to determine the optimal number of clusters ranging from two till eight. Four of the 30 indices available in NbClust demand too high computationally power when analysing a large data set, so only 26 of the indices was used. These are emphasized in appendix C. No weighing of the different indices was attempted, as this in itself is a big task perhaps more suitable for a separate study in another academic discipline such as maths and statistics. Thus, the majority rule integral in the NbClust package was therefore used as means of selection. The three alternatives with the highest score were included in further deliberations. The K-means algorithm was then run for all three alternatives deciding cluster centres and object belonging. The algorithm was run using 1000 iterations and the initial cluster centres were set to 1 in order to assure reproducibility. A summary of the script used in R is included in appendix D. The results were then exported to SPSS.

Following, the three alternative cluster solutions were evaluated in order to determine the optimal number of clusters. The objective of the cluster analysis in this study is to create segments of bicyclists that make sense in bicycle planning. This objective compares to objectives typical for market research, in a way the bicyclists are customers consuming infrastructure and the segmentation is a way of identifying different needs among the customers. Thus, it seemed relevant to employ evaluation of both face and criterion validity to aid determine the relevant number of clusters in the present study. The variables that was not included as dimensions in the cluster analysis, gender, age and experience, were used to assess criterion validity, and the criteria for face validity referenced in Mooi and Sarstedt was used for face validation. As a tool in this process ANOVA and Tukey HSD was used to assess the distinction between the clusters. Furthermore the F-value from the analysis gave an indication of how important the respective dimension was to discriminate between the clusters (StatSoft, 2013).

After determining the relevant number of clusters the cluster solution was labelled and interpreted.

5 OPERATIONALIZATION OF ANALYTICAL DESIGN

5.1 Creating the variables

5.1.1 Dependent variables

5.1.1.1 Use of different facilities

In the home questionnaire the respondents were asked how much of their bicycling the past week was on:

1. Separate bicycle path
2. Pavement
3. Bicycle lane
4. Mixed traffic

The scale of measurement was at a ratio scale with categories from 1 to 11, reflecting percentage of time 0 % till 100%. The respondents were asked to divide their total time biking in the last week, 100%, between the alternatives. This resulted in the variables being interdependent and that the distribution skewed to lower values. This was resolved by standardizing the variables before they were used in the ANOVA.

As described in chapter 2, there are three types of separate paths in Oslo, those dedicated to cyclists, those shared by cyclists and pedestrians and recreational trails. Hence it is likely that the respondents have interpreted the category in different ways, and the results should be interpreted accordingly. Another remark is that there are few stretches in Oslo that facilitates for bicyclists in mixed traffic, for example with bicycle lights and bicycle boxes, thus the mixed traffic category probably reflects time spent biking in mixed traffic without any type of facilitation for integration.

5.1.1.2 Perceived safety on different facilities

In the home questionnaire the respondents were asked how unsafe they felt:

1. Bicycling in general
2. Bicycling on a separate path shared with pedestrians
3. Bicycling in a bicycle lane
4. Bicycling in mixed traffic

The measurement scale was at ordinal level and ranged from 1 till 7, where 1 represented very safe and 7 very unsafe. The respondents could also respond 8 – "Don't know/ not an issue". The respondents who chose this last category were excluded from the analysis.

By studying how the responses on three latter questions vary from the first, we can say something about how the feeling of safety when bicycling is improved or impaired by the three designs. This presumes that the relationship on the measurement scale is linear e.g. that an improvement from 7-5 is equal to 5-3. This is not necessarily true, and this should be considered when interpreting the results.

The variables were created by subtracting the response from each of the three bottom questions from the first. This resulted in variables with positive and negative values, where positive values meant that the feeling of safety is better on the specific design than in general, and vice versa, does negative values mean that the feeling of safety is poorer on the specific design than in general. The more the values deflect from zero, the larger is the improvement or impairment of the design. Given the scale from 1-7, a deflection of ±1 equals 15 %.

5.1.2 Explanatory variables – Independent variables included in the cluster analysis

5.1.2.1 Bicycle equipment

In the home questionnaire the respondents were asked about how often they use of different types of biking equipment on utilitarian trips. The question did not refer to a specific time period. The measurement scale was at ordinal level with five categories: Always, often, sometimes, seldom and never.

The types of equipment were:

1. Helmet
2. Lights, when dark
3. Reflective vest / bicycle jacket
4. Bicycle shoes
5. Bicycle trousers
6. Heart rate monitor
7. Bicycle computer

The first two types of equipment are typically related to safety concerns. Bicyclists are recommended to wear helmets because it helps prevent head injuries (Sundfør, 2013), and lights make you visible in the dark and in bad weather. The four last types of equipment are typically for exercise purposes. The third inquiry, "reflective vest/bicycle jacket", is ambiguous. The use of reflective vest is related to visibility, and thus safety concerns, while the use of bicycle jacket can be both due to safety concerns and for exercise purposes, depending on the colour and brightness of the jacket.

In order to investigate if the use of different types of equipment actually were connected to the different latent traits suggested above, I performed a factor analysis. A factor analysis employs the variability in a selection of variables to indicate whether they represent a potentially lower number of unobservable variables e.g. factors. This is comparable to a cluster analysis, where the variability among objects is used to indicate more manageable groups. (Factor analysis is not further described in this study, and for further explanation see per example Hair, Joseph F. et al. "Ch.3: Factor analysis" in *Multivariate Data Analysis* 1998.)

The results from the factor analysis showed that the variables could be replaced by two factors. One including helmet, lights and reflective vest or jacket, and another including bicycle shoes, bicycle trousers, heart rate monitor and bicycle computer.

	Factor 1	Factor 2
Helmet		.77
Lights, when dark		.78
Reflective vest/bicycle jacket	.32	.71
Bicycle shoes	.78	.31
Bicycle trousers	.73	.39
Heart rate monitor	.79	
Bicycle computer	.77	

Note. Factor loadings < .2 are suppressed

Table 1 Factor loadings based on a principle components analysis with Varimax with Kaiser Normalisation rotation for 7 items (N= 532).

The reliability of these constructs was tested using Cronbach's Alpha. This test yields how well the variables describe the same traits, and if one or more variables should be removed in order to obtain better internal consistency (DeVellis, 2012). Helmet, lights and reflective vest or bicycle jacket yielded a result of alpha = .671 (removing reflective vest/bicycle jacket did not yield better results), and the four remaining types of equipment yielded alpha = .802. There is no universal key to what is considered a satisfactory result, but DeVellis suggests in his well cited handbook for scale development that .671 is acceptable, and .802 is very good (DeVellis, 2012).

Two new variables were constructed using the mean value from the loading variables. The scale was also turned so that low values represent little use and high values much use. With reference to the short account on possible latent traits, the two new variables were labelled *Safety equipment* and *Exercise equipment*.

Helmet is not mandatory in Norway, but it is frowned upon not to use it, and numerous attitude campaigns has been launched to increase use (Lien, 2015; Sundfør, 2013). This means that the use of safety equipment can indicate both that you are concerned with safety, and that you want to make the right choice.

5.1.2.2 Biking intensity

Also in the home questionnaire, the respondents were asked about their biking intensity on utilitarian trips. The measurement scale was at ordinal level with seven categories. 1 represented very calmly and 7 very intensively. This question was basis for the variable *Biking intensity*.

5.1.2.3 Perceived insecurity

In the home questionnaire the respondents were asked how insecure they felt bicycling on different types of infrastructure and how insecure they felt in bicycling general. The measurement scale was at ordinal level and ranged from 1 till 7. 1 represented very safe and 7 very unsafe. The respondents could also respond 8 – "Don't know/ not an issue". The respondents who chose this last category were excluded.

The question inquiring about the respondents' general level of insecurity provides insights to the respondents' perception of bicycling as an activity. This question was included as an exploratory variable in the cluster analysis and labelled *Perceived insecurity*.

5.1.2.4 Liability to break traffic regulations

In the home questionnaire the respondents are asked to consider the following statement:

"It is perfectly alright that bicyclists break some traffic regulations in order to make headway in traffic."

The measurement scale was at ordinal level ranging from 1 – Strongly agree, to 7 Strongly disagree. This provided the basis for the variable *Liability to break traffic regulations*. The variable was recoded so that high values indicated high liability and vice versa.

I interpret the indicator to mean that the bicyclists that agrees with this statement believes that the traffic regulations today is not made for or considerate enough of bicyclists and that the regulations are interfering with their ability to make headway. The reasons for agreeing with the statement can thus be due to solidarity concerns and bicycle activism, that they wish to better flow for themselves, or be solely because they feel above the law. It does not necessarily mean that they actually break the rules, but it might.

5.1.3 Explanatory variables – Additional characteristics

5.1.3.1 Gender and age

The respondents are asked about their age and gender in the field questionnaire. The variable is coded in 1 for female and 2 for male.

5.1.3.2 Experience

No question in the questionnaire is explicitly related to experience. Thus in order to operationalize this concept it was necessary to construct a new indicator.

The home questionnaire was designed to reflect the behaviour the last week, and there is only one question concerning behaviour over a longer period of time. The respondents were asked if they bike all year, and if not, when their bicycle season started and when it stops. As the respondents' bicycle season obviously have started when they are recruited to participate in the survey, the question concerning when their bicycle season started is referring to the current season. The other two questions demand that the respondents think about and answer when they normally would stop, or if they would stop at all. Thus, these questions reflect several years of behaviour. The questions are recoded to describe how many months a year the respondent bikes.

This indicator alone is not sufficient to describe the respondents' bicycle experience as we don't know anything about their biking frequency throughout the season, they might only bike a short trip every second month. Hence we need to know more about how much they bike in the months they are active. As the survey only reflect the behaviour in the past week, no such measurement exists, but we do know is how many kilometres they biked the past week. It is known that previous behaviour has a significant effect on present behaviour thus we can assume that the number of kilometres biked last week is strongly correlated with the number of kilometres biked any other week (Ajzen, 2002).

Based on the above reasoning, a variable representing experience is constructed by multiplying the number of months the respondent bikes each year and how many kilometres a week the respondent bike. The resulting variable has a considerable range, and in order to use it in the analysis it has been standardized.

5.2 Frequencies and correlations – presenting the sample

There were a total of 1075 bicyclists recruited for field questionnaire in Oslo and 581 of these agreed to answer the first home survey. After removing respondents with missing data list-wise, 532 respondents were included in the cluster analysis.

The distributions and correlations of the variables described in the previous section are presented consecutively below. These relations are referenced in the following three chapters.

The table 2 presents bivariate correlations, the mean values and standard deviations for the explanatory variables included in the cluster analysis.

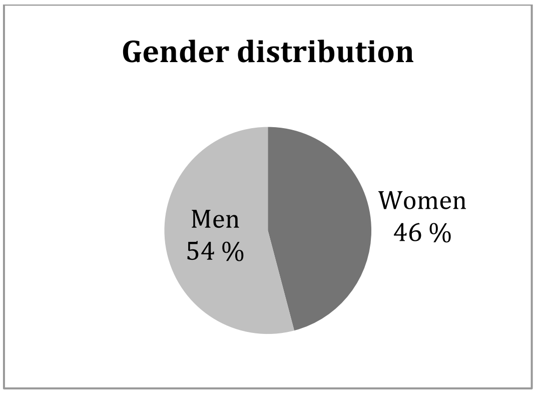
Explanatory variables[a]	1	2	3	4	5	6	7	8
1. Gender[b]	-							
2. Age	.14**	-						
3. Experience	.24**	.12**	-					
4. Unsafe bicycling	-.26**	.03	-.04	-				
5. Liability to break traffic regulations	.13**	-.10*	.05	-.14**	-			
6. Exercise equipment	.33**	.15**	.36**	-.12**	-.05	-		
7. Safety equipment	.05	.22**	.38**	.04	-.15**	.48**	-	
8. Biking intensity	.17**	-.06	.22**	-.17**	.09*	.36**	.32**	-
N	532	532	532	532	532	532	532	532
Range	1-2	12-84	2-96	1-7	1-6	1-5	1-5	1-6
M	1.54	43.81	51.63	2.83	3.13	1.98	3.80	4.26
SD	0.50	11.43	26.36	1.32	1.91	1.19	1.05	1.39

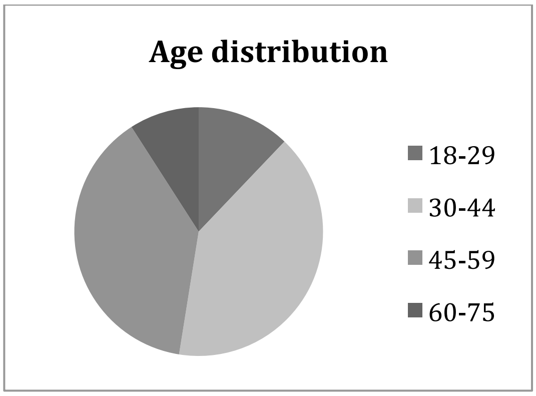
a) All variables, except b) and c) ordered 1= less, n=more b) Female=1, Male=2 *p < .05. **p < .01.

Table 2 Pearson's R, mean value, standard deviation, range and number for the explanatory variables

The correlation factors Pearson’s R, show that none of the independent variables intended for use in the cluster analysis has a correlation above .9. This means all qualify for the analysis.

The age and gender distribution in the sample is elaborated in the two sector diagrams below.





Figur 1 Age distribution in the sample

Figur 2 Age distribution in the sample

The sample has an approximately even gender distribution, but with a slightly higher share men. Approximately 80 % of the sample is between 30 and 60. The share of young adults is 12 %.

The table 3 describes the dependent variables describing frequency of use.

Dependent variables – frequency of use a)	1	2	3	4
1. Separate bicycle path	-			
2. Pavement	$-.23^{**}$	-		
3. Bike lane	$-.23^{**}$	$.11^{*}$	-	
4. Mixed traffic	$-.48^{**}$	$-.20^{**}$	$-.10^{*}$	-
M	4.18	2.86	3.46	4.73
SD	2.72	1.90	2.01	2.59
Range	1-10	1-10	1-10	1-10
N	450	459	430	501

a) All variables are ordered 1= less, n=more *p < .05. **p < .01

Table 3 Use of different bicycle facilities: Pearson correlations and descriptive statistics

Table 4 describes the average level of insecurity on different types of infrastructures. These indicators are basis for the dependent variables describing experienced change in perceived safety. These three variables are presented in the diagrams below.

Perceived unsafety a)	1	2	3
1. Unsafe bicycling on shared separate path	-		

2. Unsafe bicycling on bike lane	.39**	-	
3. Unsafe bicycling in mixed traffic	.15**	.37*	-
M	2.23	3.03	4.45
SD	1.28	1.50	1.51
Range	1-6	1-6	1-6
N	527	526	530

a) All variables are ordered 1= less, n=more *p < .05. **p < .01

Table 4 Feeling of safety: Pearson correlations and descriptive statistics

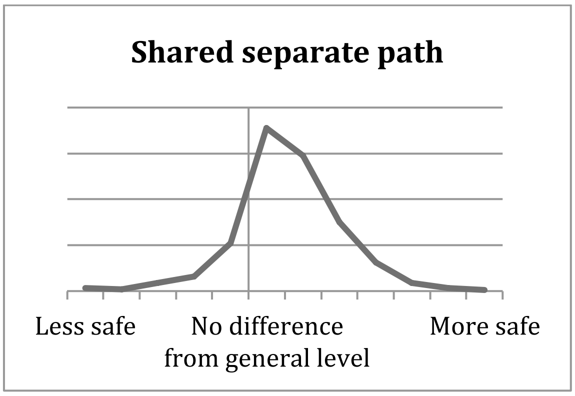


Figure 5 Distribution experienced change in perceived safety on shared separate path

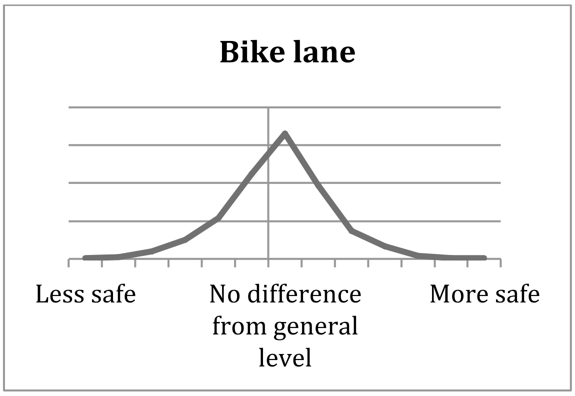


Figure 6 Distribution experienced change in perceived safety on bicycle lane

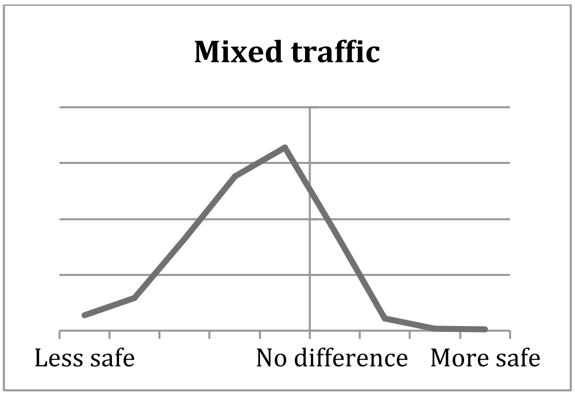


Figure 7 Distribution experienced change in perceived safety in mixed traffic

6 A BICYCLIST TYPOLOGY

This section attempts to answer the first research question. A relevant bicyclist typology is created using cluster analysis, and the results are interpreted with reference to the empirical and theoretical context of the question. As described in the methodology all variables are standardised, so I will refer to the description of sample average in the previous chapter. All clusters are evaluated relative to each other and the sample average.

6.1 The optimal number of clusters

In order to determine the optimal number of clusters in the data set, a test of 26 indices provided by the NbClust R-package was run. The test returns the number of clusters corresponding to the optimal score for each index. This is presented in the figure below.

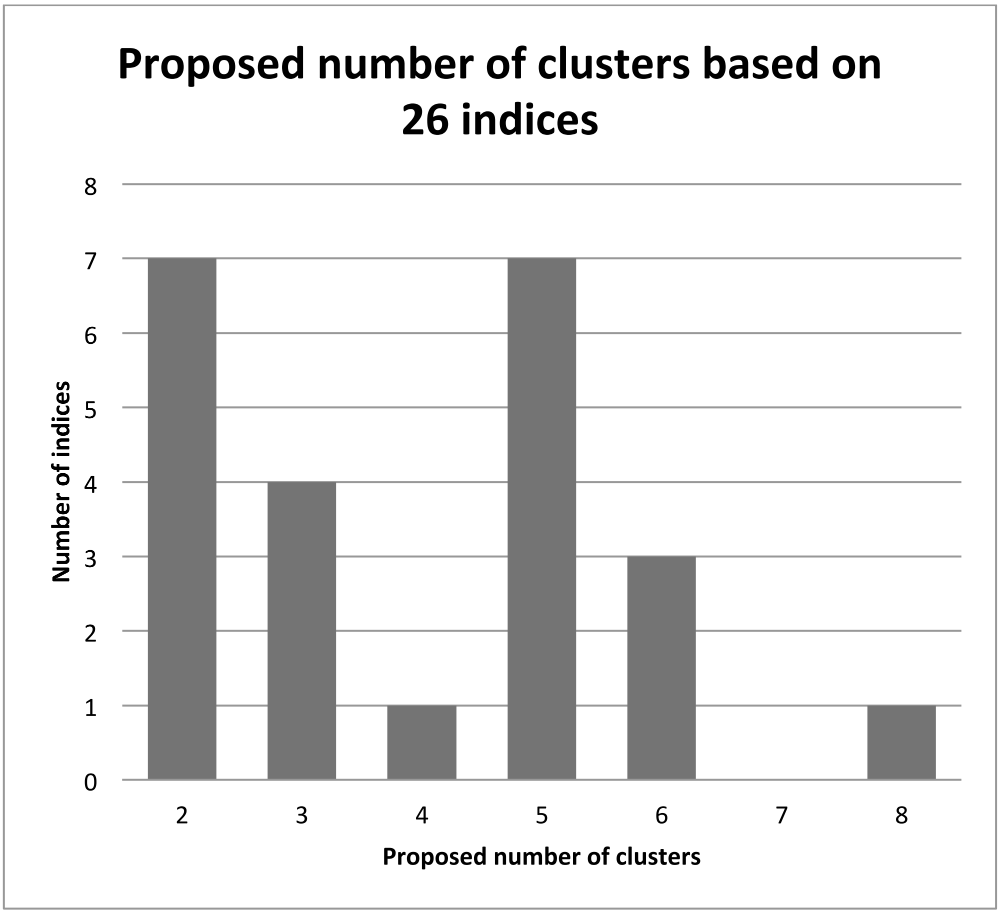


Figure 8 The proposed optimal number of clusters using the indices in the NbClust R-package

The three numbers most frequently proposed are included in further deliberations. As can be seen from the plot, two and five number of clusters is tied for first position, as they are both proposed by seven indices. Three is the third most frequently proposed number of clusters, proposed by four indices. The K-means clustering algorithm was run for each of the three proposed numbers of clusters using the input variables *Perceived insecurity*, *Liability to break traffic regulations*, *Safety equipment*, *Exercise equipment* and *Biking intensity.* The results are described in the three next sections.

6.2 Two clusters

The two cluster centres is presented in the figure below. The figure shows differences in mean value for each variable, thus illustrating what features are shared within a cluster. As mentioned earlier the variables are standardized, so mean sample value equals 0. The mean score before standardizing is presented in chapter 5.2, and should be used for reference.

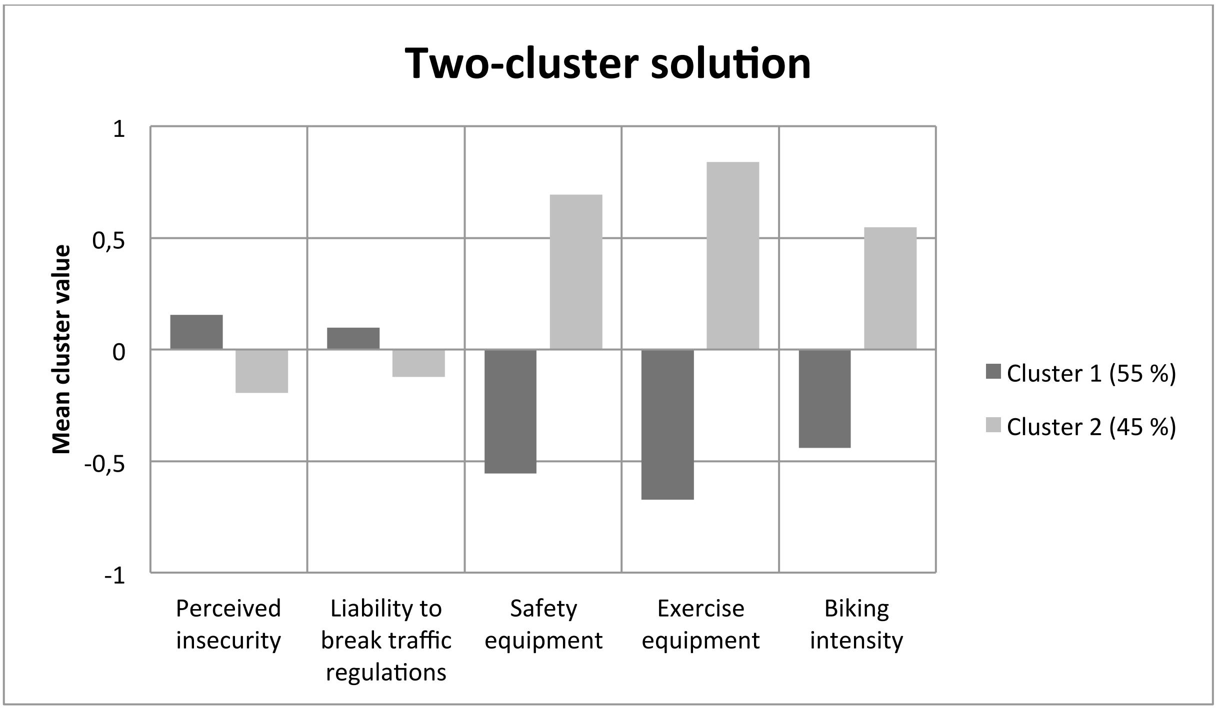


Figure 9 The two-cluster solution

The two clusters display the following characteristics:

Cluster 1: The bicyclists belonging to cluster 1 are on average a little more insecure than the average in the whole sample. They have a slightly higher liability to break traffic regulations than the average. They use less safety and exercise equipment than average and they bike less intensively.

Cluster 2: They are a little less insecure than the sample on average, and they have a slightly higher liability to break traffic regulations than the average. They use more safety and exercise equipment than average, and bike more intensively.

295 (55 %) bicyclists belong in cluster 1, and the rest of the sample, 237 bicyclists (45 %) belong in cluster 2.

An ANOVA was run to investigate in what dimensions the clusters are different from one another. It showed that there is significant difference between mean values for all variables included in the cluster analysis. This being said, the F-value for the variables *Safety equipment* and *Exercise equipment* are considerably higher than for the rest of the variables. See table below.

Variable	Perceived insecurity	Liability to break traffic regulations	Safety equipment	Exercise equipment	Biking intensity

ANOVA	F(1, 530)= 16.65, p =.00	F(1, 530)=6.48, p =.01	F(1, 530)= 332.89, p =.00	F(1, 530)= 692.09, p =.00	F(1, 530)= 68.11, p =.00

Table 5 Two-cluster solution: Results from an ANOVA testing cluster belonging against input variables in cluster analysis.

The cluster solution was analysed to see how gender, age and experience relate to cluster belonging.

The figure below shows that a larger proportion of the women in the sample are in cluster 1 than in cluster 2. A chi-square test of independence show that gender is related to cluster belonging. The relation between these variables was significant [c^2 (1, N=532) = 53.46, p = .00]. This means that women are more likely to belong to cluster 1 than cluster 2.

Figure 10 Gender distribution, age and experience in the two-cluster solution.

An ANOVA was run to investigate if there was a significant difference in mean value for age and experience between the two clusters. The test show that there is a significant difference for both age [F(1, 530)=11.42, p=.00] and experience [F(1,530)=88.91, p=.00]. As can be read from the figure above, the bicyclists in cluster 1 are slightly younger than the sample average, and are also less experienced. The bicyclists in cluster 2 are a bit older than sample average and more experienced.

6.3 Three clusters

The three cluster centres is presented in the figure below. The figure shows differences in mean value for each variable, thus illustrating what features are shared within a cluster.

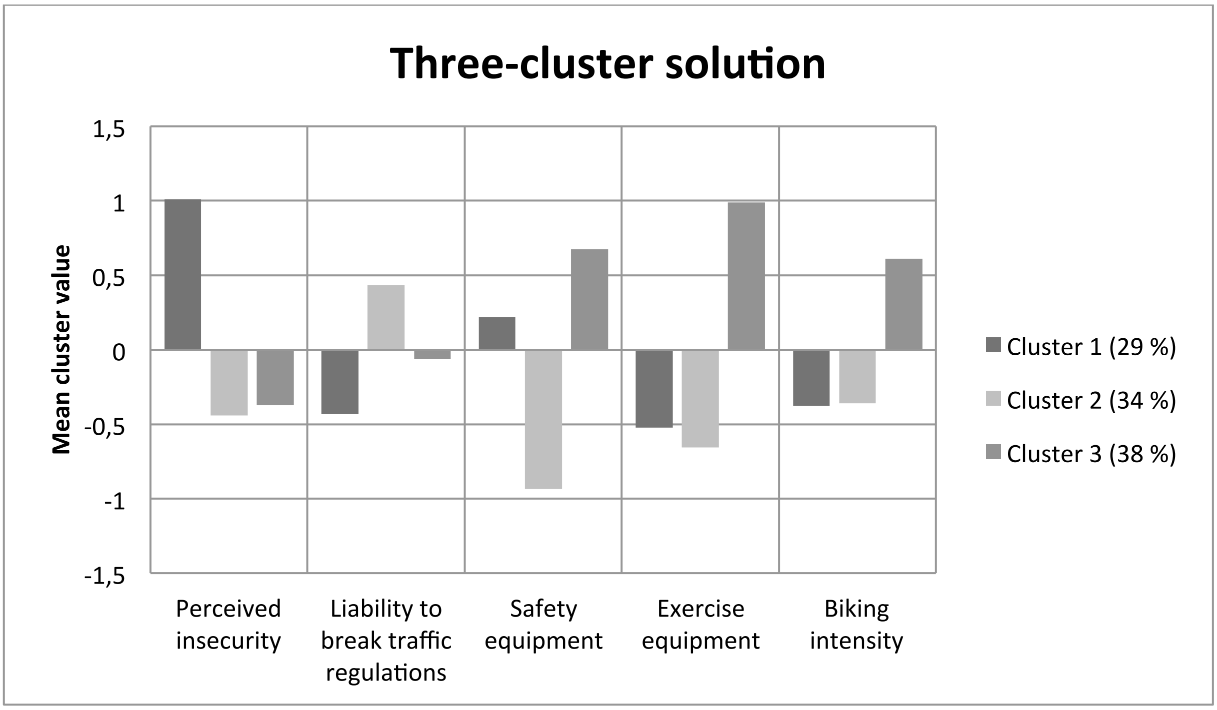


Figure 11 The three-cluster solution

The three clusters display the following characteristics:

Cluster 1: The bicyclists in cluster 1 are much more insecure than average. They are less liable to break traffic regulations than the average, and use slightly more safety equipment. They use less exercise equipment and bike less intensively than average.

Cluster 2: The most striking features within cluster 2 are that they are more liable to break traffic regulations than the average and the two other groups, and that they on average use much less safety equipment than the sample average and the two other groups. The bicyclists in cluster 2 feel more secure than average, they use less safety equipment and bike less intensively than average.

Cluster 3: The bicyclists in cluster 3 use more safety and exercise equipment than the sample average and the bicyclists belonging to the two other clusters. They also bike more intensively. They are less insecure than average, approximately equal with the bicyclists in cluster 2. They are as liable to break traffic regulations as the over all sample.

The 532 bicyclists in the sample are split so that 152 (29%) belong in cluster 1, 180 (34%) in cluster 2 and 200 (38%) belong in cluster 3.

In order to investigate in what dimensions the clusters differ from one another an ANOVA with a Tukey's HSD post hoc test was applied. The post hoc test is necessary when there are more than two groups, which is the case for this cluster solution. The test showed that there is no significant difference in mean value between cluster 2 and 3 regarding *Perceived insecurity*, neither between cluster 1 and 2 on *Exercise equipment* and *Biking intensity*. All other combination of clusters and variables showed significant difference in mean. *Liability to break traffic regulations* and *Biking intensity* display lower F-values than the other variables.

Variable	Perceived insecurity	Liability to break traffic regulations	Safety equipment	Exercise equipment	Biking intensity
ANOVA	F(2, 529)= 183.22, p=.00	F(2, 529)=35.77, p =.00	F(2, 529)= 244.49, p =.00	F(2, 529)= 384.08, p =.00	F(2, 529)=76.94, p =.00

Table 6 Three-cluster solution: Results from an ANOVA testing cluster belonging against input variables in cluster analysis.

A chi-square test of independence shows that gender is related to cluster belonging. The relation was significant [c^2 (2, N=532) = 65.23, p = .00]. Following up on this result, a post hoc test was run testing all combinations of gender and cluster belonging. This yielded that it in fact was only cluster 1 and 3 that was significantly related to gender. The gender distribution is illustrated on the figure below. It shows that cluster 1 has an overweight of women, cluster 2 is approximately evenly distributed and cluster 3 has more men than women. Thus if you belong to cluster 1 it is more likely that you are a woman, and if you belong in cluster 3 it is more likely that you are a man. Conversely if you are a woman, the distribution shows that you are only slightly more likely to belong in cluster 1 than cluster 2, and that it is not very likely that you belong in cluster 3. If you are a man you are most likely to belong in cluster 3, followed by cluster 2 and finally cluster 1. An ANOVA shows that difference in gender composition between the three clusters is significant [F(2,529)=37.0, p=.00]. Tukey's post hoc yields significant difference between all clusters.

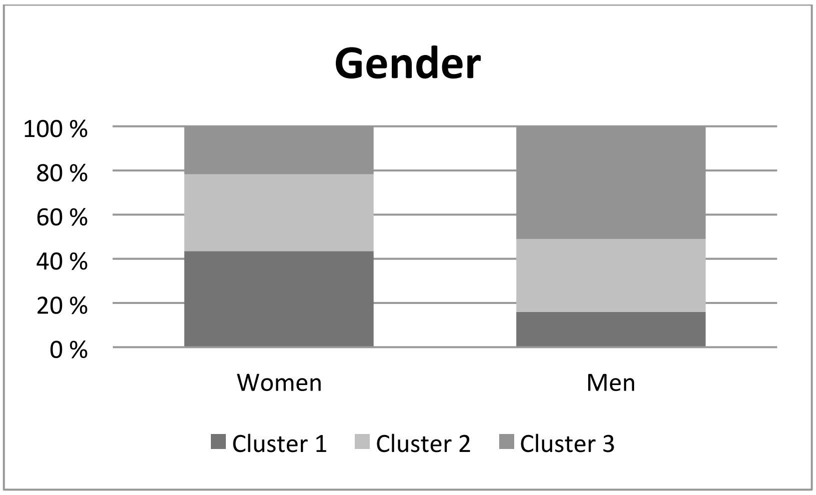


Figure 12 Gender distribution in the three-cluster solution

The descriptives, see figure 14 below, show that the bicyclists in cluster 1 are quite close to the sample average in both age and experience. The bicyclists in cluster 2 are less experienced and slightly younger, while the bicyclists in cluster 3 are more experienced and slightly older than the sample average.

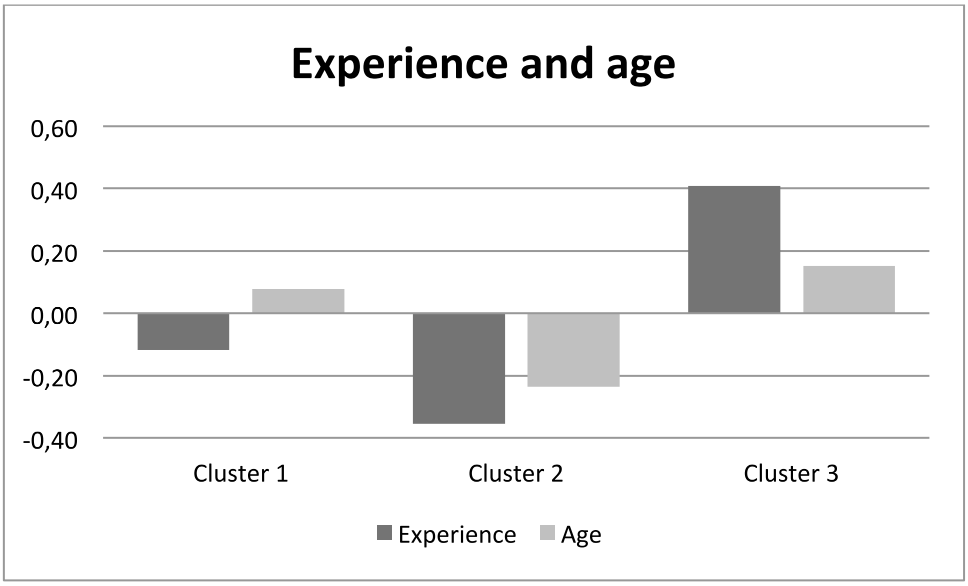


Figure 13 Differences in mean value between clusters for experience and age.

In order to explore if the differences in age and experience between the three clusters are significant, an ANOVA with Tukey's HSD was run. The ANOVA revealed that there was significant difference in age [F(2, 529)=7.94, p=.00] and experience [F(2, 529)=32.65, p=.00]. The following post hoc test exposed that there was no significant difference in age between cluster 1 and 3, and no significant difference in experience between cluster 1 and 2. All other comparisons showed significant variation.

6.4 Five clusters

The five cluster centres is presented in the figure below. The figure shows differences in mean value for each variable, thus illustrating what features are shared within a cluster.

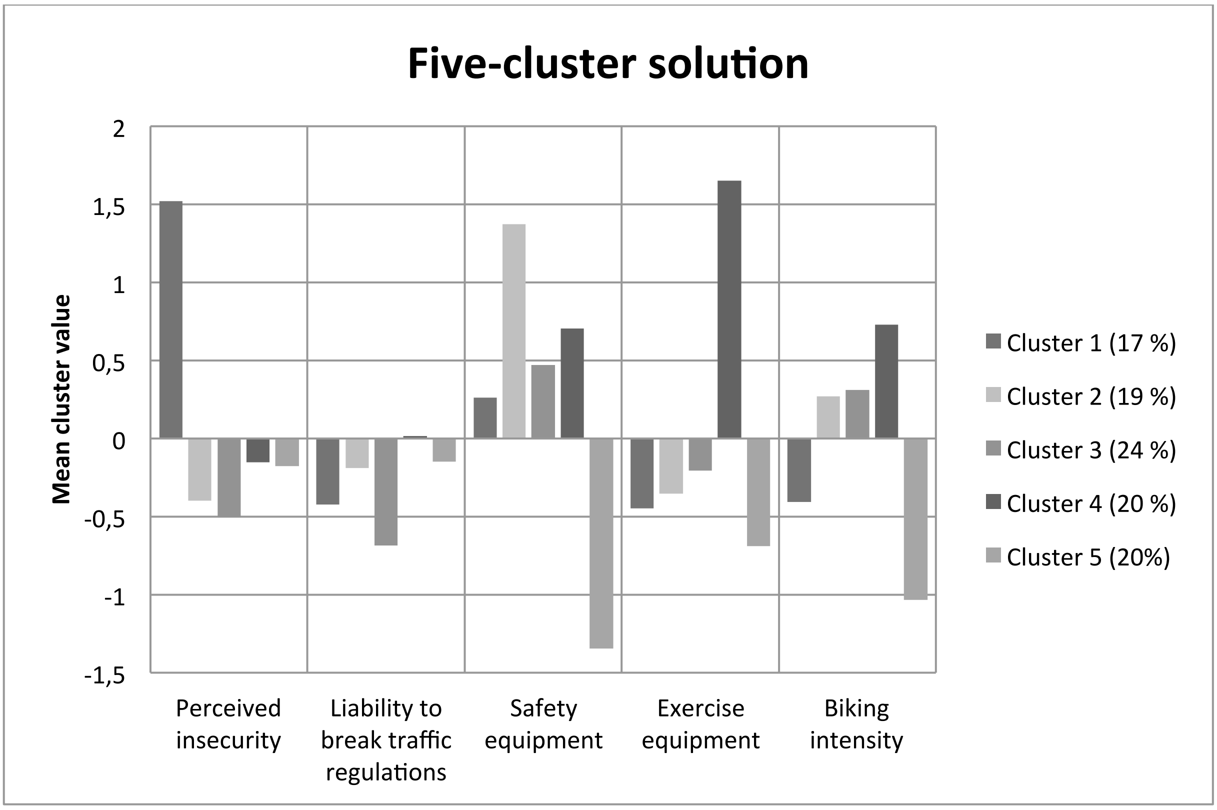


Figure 14 Five-cluster solution

The five clusters display the following characteristics:

Cluster 1: The bicyclists in cluster 1 are considerably more insecure than the sample average and all other clusters. They are less likely to break traffic regulations than average. They use less exercise equipment than average, matching their low biking intensity. The usage of safety equipment is higher than sample average, but lower than three out of the four other clusters.

Cluster 2: The most striking feature of the bicyclists in cluster 2 is that their usage of safety equipment is much higher than the sample average. The bicyclists in cluster 2 feel less insecure than average. They use less exercise equipment than average, but bike more intensely than average. Their liability to break traffic regulations is lower than average, but not more than the bicyclists in cluster 1 and 3.

Cluster 3: The bicyclists in cluster 3 feel safer than all other clusters and are also more liable to break traffic regulations. Their use of safety equipment is higher than average. The bicyclists use marginal less exercise equipment then average and bike a bit more intensely than average. In these features the bicyclists in cluster 3 are quite similar to the ones in cluster 2.

Cluster 4: The most prominent feature in cluster 4 is that biking intensity and use of exercise equipment is higher than average, and remarkably higher than among the bicyclists in the other clusters. The use of safety equipment usage is also higher than average. This cluster has a minimal higher liability to break traffic regulations than sample average, and is the only cluster to display this feature. The perceived insecurity in this group is lower than sample average, but is the second highest among the clusters.

Cluster 5: The bicyclists in cluster 5 display the decidedly lowest use of both safety and exercise equipment, as well as the lowest biking intensity. Their perception of insecurity is slightly lower than average, and so is their liability to break traffic regulations.

The distributions among the clusters are quite even, with 17% of all the bicyclists grouped in cluster 1, 19% in cluster 2, 24% in cluster 3, 20% in cluster 4 and 20% in cluster 5.

An ANOVA was applied to investigate in what dimensions the clusters differed from one another. It showed that there was significant difference in all independent variables. The following post hoc showed that not all comparisons were significant. The not significant comparisons is displayed in table 7 below.

Variable	ANOVA	Tukey's HSD: No significant difference between clusters
Perceived insecurity	$F(4, 527)=132.90, p=.00$	2 and 3; 2 and 4; 2 and 5; 4 and 5
Liability to break traffic regulations	$F(4, 527)=137.73, p=.00$	1 and 5; 3 and 4
Safety equipment	$F(4, 527)=145.84, p=.00$	1 and 3; 3 and 4

Exercise equipment	F(4, 527)=314.58, p=.00	1 and 2; 2 and 3
Biking intensity	F(4, 527)=81.15, p=.00	2 and 3

Table 7 Five-cluster solution: Results from an ANOVA and Tukey's HSD testing cluster belonging against input variables in cluster analysis.

Descriptive statistics on the gender composition within the clusters shows that cluster 1 and cluster 4 has a quite uneven gender composition. 77 % of the bicyclists in cluster 1 are women, compared to only 16 % in cluster 4. The three remaining clusters display a somewhat even composition. See figure 16 below. A chi-square test confirms that gender is related to cluster belonging [c^2 (4, N=532) = 75.16, p = .00]. The following post hoc test revealed that, as expected, only cluster 1 and 4 was related to gender. Thus women are more likely to belong to cluster 1 and men to cluster 4. An ANOVA investigating the differences in gender composition between the clusters showed that there was significant difference between clusters [F(4,527)=21.74, p=.00]. The Tukey post hoc test revealed that there was significant difference between cluster 1 and all other clusters and cluster 4 and all other clusters, but not mutually between cluster 2, 3 and 5.

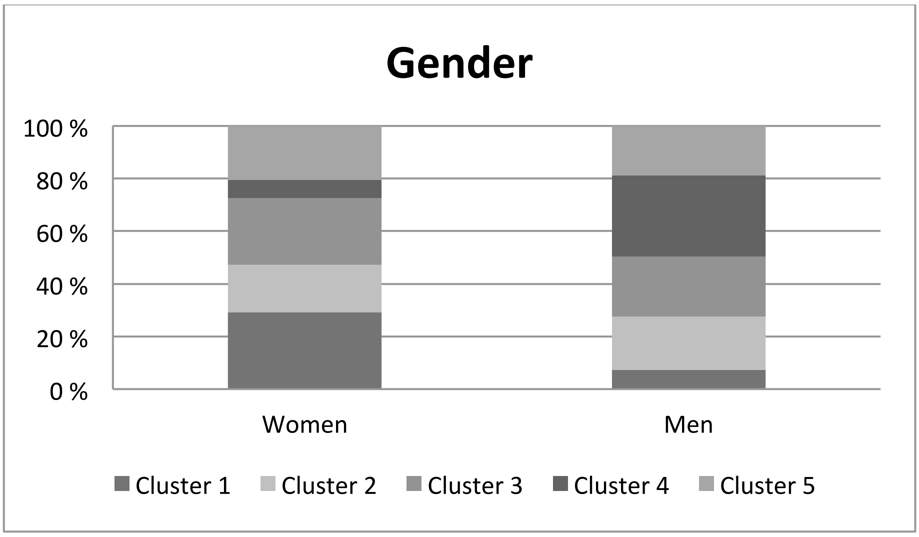


Figure 15 Gender distribution in the five-cluster solution.

The frequencies show that while the bicyclists in cluster 1, 2 and 3 seem to be fairly close to the average age and experience in the over all sample, cluster 4 are more experienced and older than average, while cluster 5 are less experienced and younger. See figure 17.

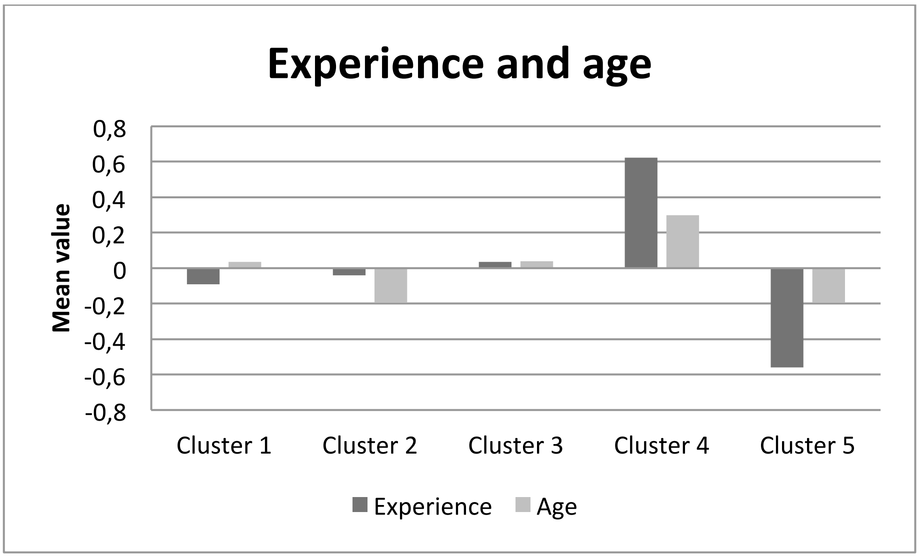


Figure 16 Mean value for experience and age for the five-cluster solution

An ANOVA with a following Tukey's HSD was run in order to explore if the bicyclist in the different clusters differ significantly from one another in age and experience. There was significant difference between the clusters for both age and experience. The results from post hoc test, however, revealed that not all clusters differed in these dimensions. The combinations of clusters where no significant relation was found are displayed in table 8.

Variables	ANOVA	Tukey's HSD: No significant difference between clusters
Experience	F(4, 527)=21.74, p=.00	1 and 2; 1 and 3; 2 and 3
Age	F(4, 527)=4.55, p=.00	1 and 2; 1 and 3; 1 and 4; 1 and 5; 2 and 3; 2 and 5; 3 and 4; 3 and 5

Table 8 Five-cluster solution: Results from an ANOVA and Tukey's HSD testing cluster belonging against experience and gender.

6.5 Determining the relevant number of clusters

The process of determining the relevant number of clusters was initiated by comparing alternative cluster solutions based on quantitative measures of validity. The three alternatives that were most frequently proposed as the optimal number, was selected to be included in further deliberations. Thus all three alternatives presented above, two, three and five clusters, assure that the final solution to will be able to expose underlying structures in the data set, and with that ensure quantitative validity. Yet, the alternatives with two and five clusters are stronger contesters in this respect as they were deemed the optimal solution by seven indices, while the alternative with three clusters was only recommended by four.

As discussed in the previous section 4.3.5.3. on how to determine the relevant number of clusters, other determinants than quantitative validity should also be evaluated. The clusters should make theoretical sense, be distinctive, recognizable, operational and informative. In order to choose between the three alternative cluster solutions these features are evaluated.

In the three previous sections the clusters solutions were explored to see how in what dimensions they differ from each other, and how they relate to theoretical related concepts e.g. age, gender and experience. This provides the basis for the following deliberation.

As can be seen from the figure 10, 12 and 15 illustrating the different cluster centres, the range of mean values increases when the number of clusters increases. This means that we get more distinctive and nuanced information when the number of clusters increases; the clusters become more “specialized”. In addition to this, can it be seen that the F-values become more even when the number of clusters increases e.g. the range in F-values decreases. This means that the five dimensions in the cluster analysis are given more equal weighting when the clusters are created. In the section about the two-cluster solution it was noted that the F-values for the variables *Safety equipment* and *Exercise equipment* were considerably higher than for the rest of the variables. While these two variables also have the highest F-values in the three- and five-cluster solution, the gap down to the remaining variables decreases, thus the dimensions’ relative influence also decreases. The effect is most detectable in the five-cluster solution, where even tough *Exercise equipment* still rank far above the other variables with 315, the remaining four rate relatively even, from 81 in *Biking intensity* to 146 in *Safety equipment*.

To sum up, a higher number of clusters allows for more nuanced information because the clusters become more specialized, and it moderates the dominating effect of certain variables. These two effects are probably closely related and due to the fact that important variation is hidden in large clusters and becomes visible when the clusters get smaller.

In the two-cluster solution there is a significant difference between the two clusters regarding *gender*, *age* and *experience*. The first cluster was dominated by women, and the bicyclists were less experienced and younger. The second was dominated by men, and the bicyclists were more experienced and older than average. These two gendered groups can also be recognised in the two subsequent solutions, but the general pattern changes somewhat. While the group dominated by men, respectively cluster 3 in the three-cluster solution and 4 in the five-cluster solution, continues to be most experienced and oldest group, the group dominated by women approaches the sample average and a new cluster assumes the characteristic of being younger and less experienced. This cluster has an approximately even gender distribution. This cluster is respectively cluster two in the three-cluster solution and five in the five-cluster solution. The changed pattern suggests that gender and experience is not inextricably linked, but is actually connected to different combinations of behaviour and perceptions. The two-cluster solution hides this story.

Based on these two aspects, will I suggest that the two-cluster solution is too simple, and fails to expose valuable information. As foreseen the solution with most clusters, the five-cluster solution, provides most information. However, it might actually provide too much. While a good solution should be informative and not hide important information, it should also be operational and recognizable.

The five-cluster solution is difficult to read. Compared to the three-cluster solution it is more difficult to imagine the type of bicyclists the different clusters represent. While a superficial interpretation of the three-cluster solution shows one cautious type, one using bicycling for exercise and one laid-back anarchist, it is difficult to give a similar offhand description of the five-cluster solution. The post hoc tests confirm this "messy situation". Few of the clusters offer unique and distinctive features that can be used in such an interpretation.

On this basis I conclude that the relevant number of clusters for this study is three.

6.6 Creating a bicyclist typology

In the previous section I gave an offhand interpretation of the three-cluster solution. In this section I will elaborate this interpretation and label the three clusters.

Cluster 1 – the Conscious Concerned: The Conscious Concerned are sensible: They follow the traffic regulations and use a helmet. They do not bike very fast, and are substantially more insecure bicycling than the other two types of cyclists: Slow and steady wins the race. The Conscious Concerned are more likely to be female than male, and they are slightly less experienced than average.

Cluster 2 – the Laid-back Rebel: The Laid-back Rebels are young and daring. They do not think much about following traffic regulations when they prohibit them from riding where they want, and are liable to break them. They bike at low intensity, doing so without any sort of bicycle equipment.

Cluster 3 - the Speedy Unafraid: The Speedy Unafraid bikes intensively for exercise purposes and their use of bicycle equipment reflects this: They always suit up before they mount their bikes. They are not insecure bicycling in Oslo and they respect the traffic regulations. The Speedy Unafraid are more likely to be male than female, and men bicycling in Oslo are most likely to be Speedy Unafraid.

7 PREFERENCE FOR INFRASTRUCTURE DESIGNS

This chapter will attempt to answer the second research question. I will investigate if the bicyclist types created in the previous chapter display difference in preference for bicycle infrastructure, and if they do what is it. This is done by looking at the difference in use of different facilities and perception of safety improvements

7.1 Frequency of use

As described in the methodology I have used an ANOVA with post hoc tests to identify significant differences in the time spent bicycling on bicycle paths, bicycle lanes, in mixed traffic and on the pavement.

The ANOVA shows that there is significant difference between the different bicyclists types regarding use of bicycle path and pavement. The effect size of this difference is 0.05 for both facilities. According to Cohen's rule of thumb this means that the effect of cluster belonging e.g. what type of bicyclist you are, on choice of infrastructure is small.

Variable	Bicycle path	Bicycle lane	Mixed traffic	Pavement
ANOVA	$F(2, 449) = 20.99$, $p=.00$	$F(2, 429) = 3.78$, $p = .15$	$F(2, 500) = 3.15$, $p = .21$	$F(2, 458) = 12.29$, $p = .00$
Effect size	eta-squared = 0.05	eta-squared = 0.01	eta-squared = 0.01	eta-squared = 0.05

Table 9 Results from an ANOVA inquiring about difference in use of infrastructure between bicyclist types.

The subsequent Tukey showed no difference between Conscious Concerned and Laid-back Rebel on neither pavement nor bicycle path, but there was significant difference between Speedy Unafraid and the two other types on both facilities.

The graph in figure 17 illustrates the difference between the types. It shows that Conscious Concerned and Laid-back Rebel use the pavement 25-30 % more than Speedy Unafraid, and that Speedy Unafraid uses bicycle paths 15-20 % more than the two other types.

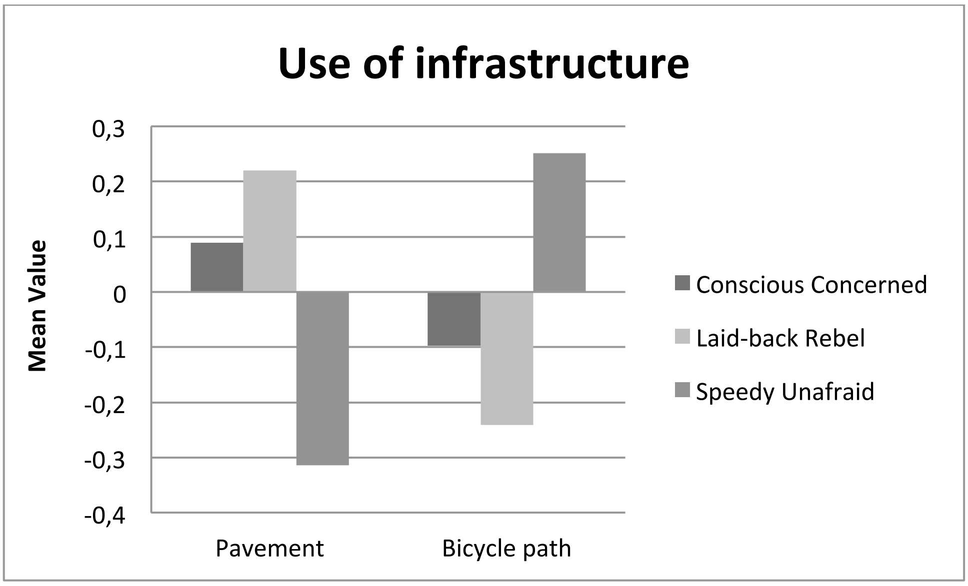


Figure 17 Difference in use of bicycle paths and pavement

7.2 Experienced change in perceived safety

I have used an ANOVA with Tukey post hoc test to investigate if the three different bicyclist types have different experiences of change in perceived safety on various infrastructure designs. The designs are mixed traffic, a separate path shared with pedestrians and bicycle lane.

Variable	Shared separate path	Bicycle lane	Mixed traffic
ANOVA	$F(2, 526) = 62.40$, $p=.00$	$F(2, 525) = 29.82$, $p =.00$	$F(2, 529) = 7.26$, $p =.00$
Effect size	eta-squared = 0.1	eta-squared = 0.19	eta-squared = 0.03

Table 10 ANOVA Experienced change in safety experience on three different infrastructures.

The ANOVA showed that there was significant difference between the bicyclist types on all infrastructures designs. As can be seen in the table below, the effect size varies between the designs. According to Cohen's rule of thumb, the effect of cluster belonging on experienced change in perception of safety on bicycle lane is of large size. The effect on experienced change in perception of safety on a shared separate path is large to medium, while it is low in mixed traffic.

The Tukey post hoc test did however reveal that there was no significant difference between Speedy Unafraid and Laid-back Rebel for any of the infrastructure designs.

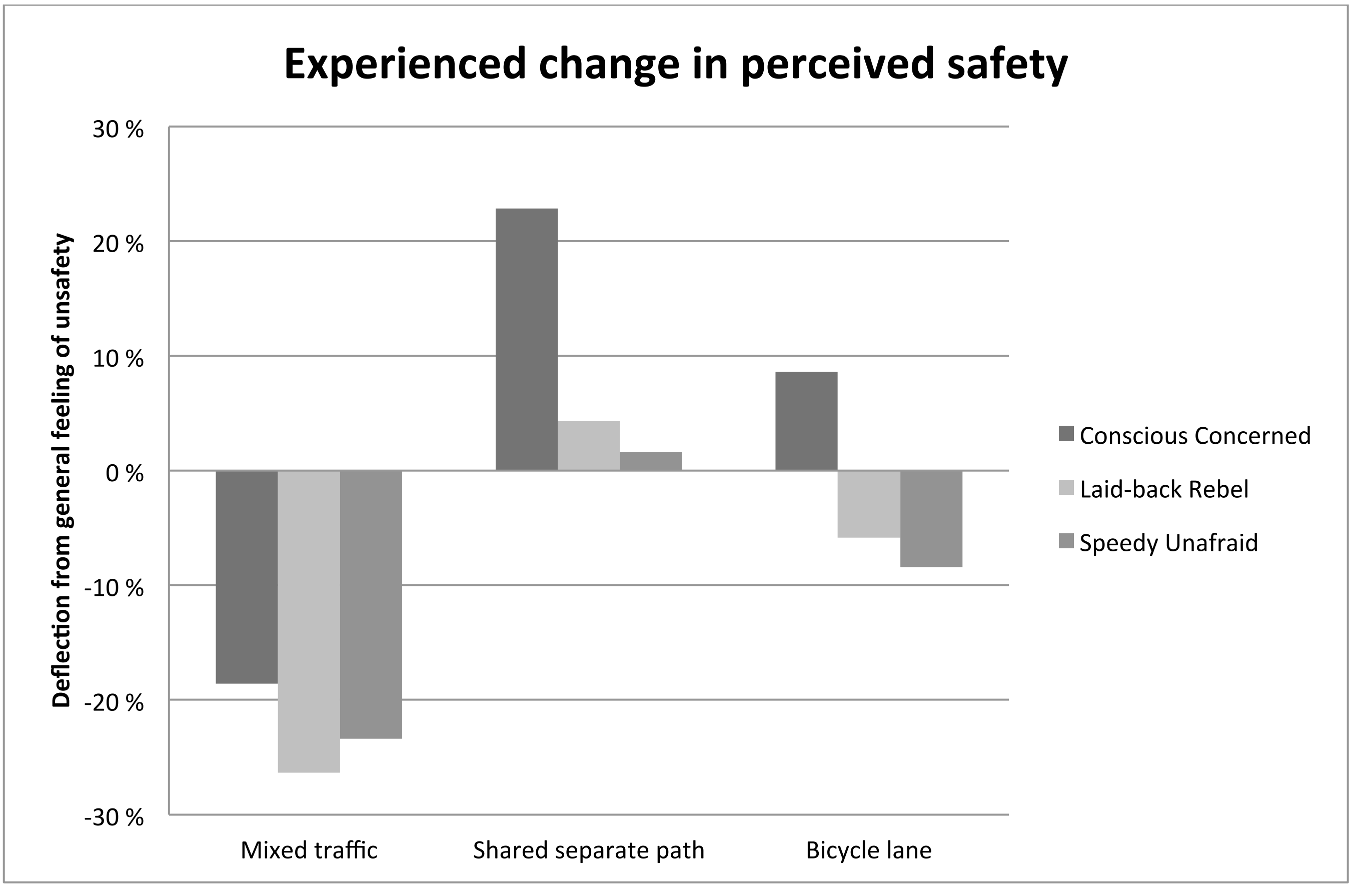


Figure 18 Experienced safety bicycling on different types of infrastructure

All types consider it substantially less safe to bicycle in mixed traffic than to bicycle in general. The deflection shows that the Conscious Concerned on average experienced a 19 % impairment of perceived safety in mixed traffic compared to 23 % for the Speedy Unafraid and 26 % in for the Laid-back Rebels.

Concerning bicycling on a separate path shared with pedestrians, all bicyclist types considered it more safe than bicycling in general. The difference between the groups were however quite large. While both Laid-back Rebels and Speedy Unafraid only experienced it marginally safer to bike on a shared path, respectively 4 % and 2 %, the average improvement for Conscious Concerned was 23 %.

Conscious Concerned on average experienced a 9 % improvement of perceived safety on bike lanes, while both Laid-back Rebels and Speedy Unafraid on average felt less safe than general on bike lanes, an impairment of respectively 6 % and 8%.

8 IMPLICATIONS FOR BICYCLE PLANNING IN OSLO

8.1 Going out of the way

On average the bicyclists spend most of their time bicycling in mixed traffic, on average 47 % of their time. The infrastructure second most used is separate bicycle paths, where the bicyclists on average spend 42 % of their time, followed by 35 % bicycle lane and 29% on the pavement. When these levels are compared with the density of connections with bicycle infrastructure in Oslo it seems evident that the bicyclists on average go out of their way to use bicycle infrastructure. This coincides with findings from other studies (Dill, 2009), and confirms that the municipality is doing right by making provisions of bicycle infrastructure an important part of the strategy to rise the bicycle share.

The bivariate correlations between time spent bicycling on different facilities suggest that time spent bicycling on other facilities is drastically reduced if the bicyclists spend a lot of their time bicycling on separate bicycle paths, while the levels are more moderate when the other types of infrastructures are measured against each other. This suggests that the connections of separate bicycle paths in some ways function as a separate parallel network, and that the bicyclists might just "go out of their way" to reach the routes with separate path designs, and follow this route as long as possible until they need to make a change of direction. This type of behaviour suggests bicycle trips with a clear purpose and of a considerable distance; it is not a short trip to the store or meeting friends near by or a stop-and-shop trip in the city. The structure of the bicycle network today supports this interpretation, as the main arteries out and in of the city centre, as well as outer ring road is facilitated predominantly using separate bicycle paths.

8.2 Three patterns

There was a significant difference in the use of pavement where Conscious Concerned and Laid-back Rebel use the pavement respectively 20 % and 30 % more than Speedy Unafraid. However, there was no significant difference between the groups concerning time spent bicycling in mixed traffic. This suggests that it is not so that they choose the same routes, but while Conscious Concerned and Laid-back Rebels choose to bike on the pavement, the Speedy Unafraid bicycle in the roadway, rather it indicates that the two constellations in general make different route choices.

The difference in use of separate bicycle paths supports this suggestion. The Speedy Unafraid use separate bicycle paths more than the other two groups, and seen in connection with the discussion above suggesting that connections with separate bicycle paths largely constitutes a parallel network, it is another evidence that Speedy Unafraid over all choose different routes than Laid-back Rebels and Conscious Concerned.

The constellation Laid-back Rebel and Conscious Concerned is however split up when looking at the difference between the types in the relative improvement of perceived safety in mixed traffic and bicycle lanes. For both Speedy Unafraid and Laid-back Rebel the impairment from their general level of safety on these two facilities is greater than for the Conscious Concerned. This means that Conscious Concerned relative experience of bicycling in mixed traffic and on bicycle lanes is better than for the two other groups, and seen in connection with their disposition as the most insecure bicyclists this suggests that Conscious Concerned on the whole seek out connections with lower traffic load than the two other types.

The relative improvement of perceived safety on shared separate path shows that the design improves the feeling of safety for all types, although the effect is considerably larger for Conscious Concerned than the two other types. This supports the suggestion made earlier that the Conscious Concerned bike mostly in their immediate environment, and that seeking out the connection with this type of facilitation is not worth it for their trip purposes.

Summing up, the above indications delineate three type-specific patterns. Conscious Concerned seek out connection with lower traffic load. Given their needs it is not worth it to go out of their way to seek out separate bicycle path, even though they perceive it as the relatively best facility. Similar to Conscious Concerned, Laid-back Rebel bicycles in the immediate environment. They do not bother to seek out routes with low traffic load. They just choose the most convenient route. The Speedy Unafraid goes out of their way to find the best connections.

A hesitation concerning the results derived concerning safety perception on different designs is that the variables are constructed to illustrate deflection from the bicyclists' general perception of safety bicycling in Oslo. This perception could be influenced by where the bicyclists normally bike, meaning that if a bicyclist normally bike on one type of design, then their general perception of safety bicycling could reflect the experience of bicycling on that particular design. The tendencies for Speedy Unafraid could suggest that this effect is present, as they bike more than the others on separate bicycle path, and only experience a marginal improvement of perceived safety on the design. However the improvement is equally marginal for Laid-back Rebel, and they have bike considerably less on separate bicycle paths than Speedy Unafraid and the over all average. Thus there can be made no clear conclusions about the effect

8.3 Planning for all types

Based on the above analysis of different bicyclist types and preferences for infrastructure designs I suggest the following recommendations in order to plan for the different types needs and travel patterns:

Conscious Concerned: This bicyclist type need bicycle infrastructure that is not "out of their way", the facilitations have to be in the immediate environment. This means that they would benefit from a denser bicycle network with low mask width. The infrastructure design should preferably be separate bicycle paths.

Laid-back Rebel: Similar to Conscious Concerned the Laid-back Rebels need bicycle infrastructure in the immediate environment. However they do benefit much from a separate design, nor bicycle lane. A prominent characteristic with the Laid-back Rebels is that they are prone to break traffic regulations in order to make headway. This indicates the best type of facilitation is interventions that make it easier for them to manoeuvre in mixed traffic. Examples of such are bicycle lighting and bicycle boxes.

Speedy Unafraid: This type bike for exercise, and their travel pattern suggests that they are willing to go out of their way to seek out appropriate facilitation for a direct and convenient route. Thus a bicycle expressway seems like a suitable type of facilitation.

These recommendations testifies to the fact that the findings from this study far a long supports the direction of infrastructure provision planning delineated in the bicycle strategy and the consultation draft for the new bicycle network in Oslo.

In the Spacescape evaluation of bicycling in Oslo, they describe the typical Oslo bicyclist, as a fearless male. This prototype was also basis for the first bicyclist profile; *Those who bikes no matter what*, used in a focus group to evaluate different bicycle standards. According to the typology featured in this study, this description is valid for two types, namely the Speedy Unafraid and the Laid-back Rebels. There was no difference in the level of insecurity between the two groups and they were both more confident than the sample average. It is slightly more likely that a man belongs to Speedy Unafraid than Laid-back Rebel, but not by that much. This means that the typical Oslo bicyclist is actually, more accurately described as two types, with distinctively different behavioural and travel patterns. If similar deliberations on bicycle standard are initiated in connection with the formulation of the Oslo standard, Oslo municipality and the Norwegian Public Road Administration would benefit from inviting bicyclists representing both types.

9 DISCUSSION

The purpose of this section is to discuss the relevance of the findings in this study. I will discuss what considerations should be made when employing travel behaviour and cluster methodology to investigate mobility for planning purposes.

9.1 Including non-bicyclists

This study only includes people that already bike in Oslo. In order to recruit new bicyclists it would be a good idea to also include those who do not bike in such as preference study in order to see what types of designs or infrastructure provisions that might get them to bike. Other studies have contemplated similar concerns. Hunt and Abrahams (2007) argues in their earlier described study of infrastructure preferences that asking non-cyclists has been criticised because they have no experience using the infrastructure, and thus does not know first hand what they prefer (2007). Continuing along the lines of such an argumentation, one could say that those who do not bicycle themselves form a perception about the infrastructure designs based on narratives formulated by those who have experienced it for themselves, observations in traffic or the media's representation. Asking non-bicyclists will thus be second-hand information, and then you might as well go straight to the source.

Conferring the theory of planned behaviour this attitude is challenged. If the beliefs about an activity have formed, it does not matter if it derived from own experiences or not, and it does not matter if it is right or wrong. The belief determines behaviour. So if the aim is to get non-cyclists to bike, then their input is valuable. However, non-cyclists will become bicyclists when they start to bike. Given the fact that the bicycle share in Oslo has increased over the years it seems plausible that some of the respondents with little experience were recently non-cyclists. If we assume that there is a continuity concerning the diversity of new bicyclist recruits, the preferences of those who currently are non-cyclists will soon become the preferences of those with little experience.

9.2 Relying on the individual - the planner's responsibilities

This study was based on the hypothesis that studying bicyclists' individual perceptions is an important tool in order to increase the bicycle share in cities. However, taking point of departure in individuals' perceptions means that the planner interpreting the results and employing them for planning purposes, needs to be aware that there might be a discrepancy between perceptions of an activity and the actual activity. For bicycling this is particularly relevant concerning perceived and actual safety.

Being separated from motorised traffic is often perceived by bicyclists to be a safer and more comfortable option than mingling (Forsyth & Krizek, 2010; Walker, 2011). This is also true for the population in this study, as can be seen in the distributions of improvement of perceived safety on the different facilities in section 5.2. Yet, there is little evidence to suggest segregated infrastructure actually is safer than the integrated approach (Forsyth & Krizek, 2010; Walker, 2011). Actually, most conflicts occur in intersections, and not when two modes are travelling in the same direction. As segregated infrastructure designs often make intersections more complicated, several professionals regard the integrated approach to be the safer option (Forsyth & Krizek, 2010). This signifies that perceived safety and actual safety is not necessarily coincided, and that many consider choosing between segregated and integrated infrastructure equal to choosing between perceived and actual safety.

The Norwegian Public Roads Administration (NPRA) has earlier stated that they are sceptical of Oslo municipality's desire to build segregated infrastructure precisely for this reason, but is interested in a dialogue on the subject (Yset, Martinsen, & Hansen, 2014). The proposals for the new Oslo-standard in the consultation draft for the new bicycle network, suggest however that the NPRA might have changed their opinion (Statens Vegvesen & Oslo Kommune, 2015).

Another aspect the planner has to consider taking point of departure in individual perceptions is that large-scale mechanisms are not captured by an aggregation of individual preferences. The inherent logic in the utility maximizing approach to travel behaviour is that each individual acts to maximize its own utility, however these choices might not be efficient in order to achieve welfare for a group as a whole (Sen, 1973). An example of this is that if everybody bicycled in the middle of street, which is allowed according to Norwegian legislation, the cars would need to adapt to the situation and slow down.

The two aspects described in this section aim to describe that a planning decision should not be based on an experiment taking point of departure in individuals' expressed preference alone. The planner has a responsibility to qualify the results by looking at discrepancies between perception and reality, and investigating the possibility for the existence of large scale mechanisms that could affect preference or the conditions in the longer term.

9.3 Explaining mobilities?

The mobility turn is an alternative to the mainstream approach to travel behaviour described in chapter 3. It is based on the recognition that there is a need for new theoretical frameworks and development of methods that can capture the deeper understanding of mobilities, meaning what takes place between two points in space, the situations and considerations made when you are on the move (Sheller & Urry, 2006). In the traditional travel behaviour literature travelling has been treated as a 'black box', perceived as a value-neutral process derived from rational demand and the focus has been solely on how or why people move in space (Spinney, 2009). The mobility turn suggests that we need mobile methods to understand the meaning of mobility and have proposed several approaches to do so. Two examples of such methods that could be relevant to get a better understanding of bicycling on different type of infrastructure designs are (1) observations of bicyclists, studying their pattern of behaviour, timing and interaction with others traffic users and physical barriers, and (2) ethnographical studies and "ride-along" studies bicyclists on the move (Spinney, 2009).

This study might have benefited from being supplemented by methodology from the mobility turn. A ride-along session with one or more bicyclists from each of the different types could give insights to the causal relations determining preferences and choice of route. It would also be possible to address specific details or concerns more intimately. Such an approach would give more nuanced and rich results. On the other hand, such a method would imply that you study fewer bicyclists and that the results are more difficult to generalise from.

10 CONCLUSION

This study aims to explore what different bicyclist types we have in Oslo, if these types display preference for different type of infrastructure designs and what implications this has for bicycling in Oslo.

A cluster analysis was made including different beliefs or perceptions about bicycling as dimensions. The clusters were compared based on gender distribution and experience in order to increase their explanatory power and make them relatable to previous studies. The cluster analysis suggested that we have three types of bicyclists in Oslo; the Conscious Concerned, the Laid-back Rebel and the Speedy Unafraid.

Type belonging was used as an independent variable to investigate if it could explain difference in time spent bicycling on separate bicycle path, in mixed traffic, on the pavement and in bicycle lanes, and if it could explain difference in perception of relative safety improvement bicycling on bike lanes, separate bicycle paths and in mixed traffic, assuming that these indicators together could say something of difference in preference between the groups.

Out of the three types it seems that providing new bicycle infrastructure, especially separate bicycle paths, brings the largest improvement of perceived safety to Conscious Concerned. The Speedy Unafraid uses separate bicycle paths more than the other two types, and Conscious Concerned and Laid-back Rebel bike more on the pavement than Speedy Unafraid. Given the spatial structure of the routes with this separate bicycle path design, it seems that the general travel pattern of Speedy Unafraid differs from the two other types. Speedy Unafraid seek out separate bicycle paths, while Conscious Concerned and Laid-back Rebel bike in the immediate environment, explaining much use of the pavement. The interpretation suggesting different travel patterns is supported by the difference between the types in the relative improvement of perceived safety in mixed traffic and bicycle lanes. Both Speedy Unafraid and Laid-back Rebel considers bicycling in mixed traffic and on bicycle lanes worse than Conscious Concerned, suggesting that Conscious Concerned deliberately chooses connections with lower traffic load than the two other types, while Speedy Unafraid might choose the route with the best flow.

The findings suggest that the Conscious Concerned prefer to bike in the immediate environment, but choose the routes with low traffic load, the Speedy Unafraid seek out routes with good flow; separate bicycle paths or large connections, and can go “out of their way” to reach these connections, and the Laid-back Rebel bikes in the immediate environment, but does not necessarily bother to seek out the routes with low traffic load. A denser bicycle network is an essential provision in order to facilitate for bicycling for the first two types, while a bicycle expressway would benefit the Speedy Unafraid.

The findings resonate with the existing literature on preference for different bicycle environments. The travel pattern of Speedy Unafraid coincides with the finding that experienced bicyclists go out of their way to find good connections (Dill, 2009). Previous findings suggest that women prefer segregated designs (Garrard et al., 2008; Monsere et al., 2012), this coincides with the fact the Conscious Concerned who are more likely to be female, benefit the most from bicycling on separate bicycle paths. Stinson and Bhat (2005) found that experienced cyclists are less concerned with the level of separation from motorised transport than inexperienced cyclists, which concurs with the choices of the more experienced Speedy Unafraid.

This study illustrates the relevance of using cluster analysis for planning purposes. However, taking point of departure in individual perceptions means that the planner interpreting the results and employing them for planning purposes needs to be aware that there might be a discrepancy between perceptions of an activity and the actual activity. For bicycling this is particularly relevant concerning perceived and actual safety. The planner also has to consider large-scale mechanisms that are not captured by an aggregation of individual perceptions.

Dealing with urban mobility, the quantitative approach featured in this study, could benefit from being supplemented with methodologies from the mobility turn allowing further explorations of causal relations. A desirable approach could be a combination of both where a cluster analysis form basis for the analysis by suggesting relations that should be investigated further. Hence, the Conscious Concerned, the Laid-back Rebel and the Speedy Unafraid, might be helpful in coming studies of bicycle infrastructure when planning for Oslo as a bicycle city for the future.

REFERENCES

Ajzen, I. (1991). The theory of planned behavior. *Organizational behavior and human decision processes, 50*(2), 179-211.

Ajzen, I. (2002). Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and social psychology review, 6*(2), 107-122.

Bakeman, R., & Robinson, B. F. (2005). *Understanding statistics in the behavioral sciences*: Psychology Press.

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice, 46*(10), 1730-1740. doi:10.1016/j.tra.2012.07.005

Bryman, A. (2012). *Social research methods*: Oxford university press.

Bymiljøetaten. (2015). Sykkelveinettet i Oslo. Retrieved from http://www.arcgis.com/home/webmap/viewer.html?webmap=15f55d0441294bfca8b2aae44ca8b7e7&extent=10.7009,59.9061,10.8276,59.9483

Caulfield, B., Brick, E., & McCarthy, O. T. (2012). Determining bicycle infrastructure preferences – A case study of Dublin. *Transportation Research Part D: Transport and Environment, 17*(5), 413-417. doi:10.1016/j.trd.2012.04.001

Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: an R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, 61*(6), 1-36.

City of Johannesburg. (2009). *Framework for Non-Motorised Transport*. Retrieved from Johannesburg: http://www.joburg-archive.co.za/2009/pdfs/transport/nmt_framework09.pdf

Cleary, J., & McClintock, H. (2000). Evaluation of the Cycle Challenge project: a case study of the Nottingham Cycle-Friendly Employers' project. *Transport Policy, 7*(2), 117-125.

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences (rev*: Lawrence Erlbaum Associates, Inc.

Damant-Sirois, G., Grimsrud, M., & El-Geneidy, A. M. (2014). What's your type: a multidimensional cyclist typology. *Transportation, 41*(6), 1153-1169.

Dennis, K., & Urry, J. (2009). *After the car*: Polity.

DeVellis, R. F. (2012). *Scale development: Theory and applications* (Vol. 26): Sage publications.

Dill, J. (2009). Bicycling for transportation and health: the role of infrastructure. *Journal of public health policy*. doi:10.1057/jphp.2008.56

Forskningsrådet. (2014). Sikkerhet for myke trafikanter i bytrafikken. In Forskningsrådet (Ed.), *Forskningsrådet*. Oslo: Norges forskningsråd

Transportsikkerhet – TRANSIKK

www.forskningsradet.no/transikk.

Forskrift om kjørende og gående trafikk (trafikkregler), (1986).

Forsyth, A., & Krizek, K. J. (2010). Promoting walking and bicycling: assessing the evidence to assist planners. *Built Environment, 36*(4), 429-446.

Fyhri, A., & Sundfør, H. B. (2013). *Arbeidsdokument 50543: Datainnsamlig og dataprepararing - Safety in Numbers*. TØI. Oslo.

Garrard, J., Rose, G., & Lo, S. K. (2008). Promoting transportation cycling for women: the role of bicycle infrastructure. *Prev Med, 46*(1), 55-59. doi:10.1016/j.ypmed.2007.07.010

Halkidi, M., Batistakis, Y., & Vazirgiannis, M. (2001). On clustering validation techniques. *Journal of Intelligent Information Systems, 17*(2), 107-145.

Handy, S. (2005). Critical assessment of the literature on the relationships among transportation, land use, and physical activity. *Transportation Research Board and the Institute of ....* Retrieved from http://onlinepubs.trb.org/onlinepubs/archive/downloads/sr282papers/sr282handy.pdf

Handy, S., Cao, X., & Mokhtarian, P. (2005). Correlation or causality between the built environment and travel behavior? Evidence from Northern California. *Transportation Research Part D: Transport and Environment, 10*(6), 427-444.

Handy, S., Van Wee, B., & Kroesen, M. (2014). Promoting cycling for transport: Research needs and challenges. *Transport Reviews, 34*(1), 4-24.

Hjorthol, R. E., Øystein; Uteng, Tanu Priya. (2014). *Den nasjonale reisevaneundersøkelsen 2013/14 - nøkkelrapport*. Retrieved from Oslo:

Hunt, J. D., & Abraham, J. (2007). Influences on bicycle use. *Transportation, 34*(4), 453-470.

Jacobsen, P. L. (2003). Safety in numbers: more walkers and bicyclists, safer walking and bicycling. *Injury prevention, 9*(3), 205-209.

Julien, A. (2000). Comparaison des principaux manuels européens d'aménagements cyclables: aménagement cyclable et espace urbain.

Julsrud, T. E. (2012). *Everyday travel patterns* (1191/2012). Retrieved from Oslo:

Knox, P., & Pinch, S. (2014). *Urban social geography: an introduction*: Routledge.

Landau, S., & Ster, I. C. (2010). Cluster analysis: overview. *International Encyclopedia of Education, 3*, 72-83.

Lien, M. (2015). Sykkelhjelm - en norsk helligdom. *Morgenbladet*. Retrieved from http://morgenbladet.no/2015/09/sykkelhjelm-en-norsk-helligdom

Mayor of London. (2015). Mayor's vision for cycling. Retrieved from https://www.london.gov.uk/priorities/transport/cycling-revolution

McClintock, H. (1992). *Bicycle and City Traffic*: Belhaven Press.

McFadden, D. L. (2002). The path to discrete-choice models. *ACCESS Magazine, 1*(20).

Monsere, C. M., McNeil, N., & Dill, J. (2012). Multiuser Perspectives on Separated, On-Street Bicycle Infrastructure. *Transportation Research Record: Journal of the Transportation Research Board, 2314*(-1), 22-30. doi:10.3141/2314-04

Mooi, E., & Sarstedt, M. (2011). 9 Cluster Analysis *A concise guide to market research: The process, data, and methods using IBM SPSS statistics*: Springer Science & Business Media.

Pucher, J., Dill, J., & Handy, S. (2010). Infrastructure, programs, and policies to increase bicycling: An international review. *Preventive Medicine, 50, Supplement*(0), S106-S125. doi:http://dx.doi.org/10.1016/j.ypmed.2009.07.028

Robinson, D. L. (2005). Safety in numbers in Australia: more walkers and bicyclists, safer walking and bicycling. *Health promotion journal of Australia, 16*(1), 47-51.

Samferdselsdepartementet. (2013). *Meld. St. 26 Nasjonal transportplan 2014-2023*. (Stortingsmelding 26). Oslo.

Sen, A. (1973). Behaviour and the concept of preference. *Economica*, 241-259.

Sheller, M., & Urry, J. (2006). The new mobilities paradigm. *Environment and Planning A, 38*, 207-226.

Sims, R., Schaeffer, R., Creutzig, F., Cruz-Núñez, X., D'Agosto, M., Dimitriu, D., . . . Tiwari, G. (2014). Transport. In O. Edenhofer, R. Pichs-Madruga, Y. Sokona, E. Farahani, S. Kadner, K. Seyboth, A. Adler, I. Baum, S. Brunner, P. Eickemeier, B. Kriemann, J. Savolainen, S. Schlömer, C. von Stechow, T. Zwickel, J.C. Minx (Ed.), *Climate Change 2014: Mitigation of Climate Change. Contribution of Working Group III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge, United Kindom and New York, NY, USA: Cambridge University Press.

Singapore Government. (2015). Cycling. Retrieved from http://www.mot.gov.sg/About-MOT/Land-Transport/Cycling/

Spacescape. (2014a). *Evaluering av Oslos sykkelstrategi 2005-2015. Underlagsrapport for sykkelstrategi for Oslo.* Retrieved from Oslo:

Spacescape. (2014b). *Kartlegging av Oslosyklisten*. Retrieved from Oslo:

Spacescape. (2014c). *Sykkelstrategi 2015-2025*. Retrieved from Oslo:

Spinney, J. (2009). Cycling the city: Movement, meaning and method. *Geography Compass, 3*(2), 817-835.

Statens Vegvesen. (2013). *Sykkelhåndboka. Utforming av sykkelanlegg. Veiledning. Håndbok V122.* Retrieved from http://www.vegvesen.no/_attachment/69912/binary/964012?fast_title=H%C3%A5ndbok+V122+Sykkelh%C3%A5ndboka.pdf

Statens Vegvesen & Oslo Kommune. (2015). *Plan for sykkelveinettet i Oslo. Høringsutgave*. Retrieved from Oslo:

StatSoft, I. (2013). *Electronic Statistics Textbook* Retrieved from http://www.statsoft.com/textbook/ Retrieved from http://www.statsoft.com/textbook/

Stinson, M. A., & Bhat, C. R. (2003). Commuter bicyclist route choice: analysis using a stated preference survey. *Transportation Research Record: Journal of the Transportation Research Board, 1828*(1), 107-115.
Stinson, M. A., & Bhat, C. R. (2005). A comparison of the route preferences of experienced and inexperienced bicycle commuters. *Transportation Research Board Paper*(05-1434).
Sundfør, H. B. (2013). *Et trygt valg eller en god vane? Hvordan forklare bruk av sykkelhjelm med en utvidet modell for overveid handling.* NMBU, Ås.
The City of New York. (2014). Bicycle Network Development. Retrieved from http://www.nyc.gov/html/dcp/html/bike/home.shtml
Tilahun, N. Y., Levinson, D. M., & Krizek, K. J. (2007). Trails, lanes, or traffic: Valuing bicycle facilities with an adaptive stated preference survey. *Transportation Research Part A: Policy and Practice, 41*(4), 287-301. doi:http://dx.doi.org/10.1016/j.tra.2006.09.007
UN-Habitat. (2013). *State of the world's cities 2012/2013: Prosperity of cities*: Routledge.
United Nations. (1992). *Agenda 21: Programme of Action for Sustainable Development*. Retrieved from New York:
United Nations. (2014). *World Urbanization Prospects: The 2014 Revision, Highlights (ST/ESA/SER.A/352)*. Retrieved from
Urry, J. (2007). *Mobilities*: Polity.
Walker, I. (2011). Bicyclists. In B. E. Porter (Ed.), *Handbook of traffic psychology*: Academic Press.
Wardman, M. (1988). A comparison of revealed preference and stated preference models of travel behaviour. *Journal of Transport Economics and Policy*, 71-91.
Yset, S. S., Martinsen, E., & Hansen, A. H. (2014, 11.09.2014). Nå kan Oslo få egne sykkelveier. *nrk.no*. Retrieved from http://www.nrk.no/ostlandssendingen/vil-ha-adskilte-sykkelveier-i-oslo-1.11924608

APPENDICES

Appendix A – Home questionnaire

ID:intro

Information
Velkommen til skjemaet Vi vil gjerne vite litt mer om hvordan du har opplevd det å sykle i den siste uka. Det vil si i går, og de seks dagene forut.

starttid	Tid for oppstart av intervjuet
◆ range:* ◆ afilla:sys_timenowf c Fylles inn automatisk	☐☐☐☐☐☐ 1

startdato	Dato for oppstart av intervjuet
◆ range:* ◆ afilla:sys_date c Fylles inn automatisk	☐☐☐☐☐☐☐☐ 1

respondent	Respondentnummer til matching
◆ afilla:sms_Respid c Genereres automatisk	☐☐☐☐☐☐☐☐☐ 1

SykletIUka	Har du syklet i løpet av den siste uka?	
◆ range:*		
Ja		○ 1
Nei		○ 2

Information
◆ exit:yes ◆ filter:\SykletIUka.a=2 ◆ redirect:http://www.toi.no Siden du ikke har syklet i løpet av den siste uka, er du desverre ikke i målgruppen for undersøkelsen. Takk for at du tok deg tid til å svare.

SykkelMiljo	Tenk på dine sykkelturer den siste uka, og på hvordan de fordeler seg på ulike typer steder. Hvor mye av din sykling har foregått på										
	0%					50 %					100 %
	1	2	3	4	5	6	7	8	9	10	11
Fortau	○	○	○	○	○	○	○	○	○	○	○ 1
På separat sykkelvei	○	○	○	○	○	○	○	○	○	○	○ 2
På sykkelfelt i gate/vegbane	○	○	○	○	○	○	○	○	○	○	○ 3

SykkelMiljo	Tenk på dine sykkelturer den siste uka, og på hvordan de fordeler seg på ulike typer steder. Hvor mye av din sykling har foregått på											
I gate/veg uten sykkelfelt	○	○	○	○	○	○	○	○	○	○	○	4
I skog og mark etc.	○	○	○	○	○	○	○	○	○	○	○	5

Information
Nå kommer noen spørsmål om hvordan det har vært å sykle i den siste uka, og ditt samspill med andre trafikanter. Først vil vi at du tenker tilbake på hvordan du har opplevd å møte BILISTER i ulike situasjoner (kryss, rundkjøringer, avkjørsler, parkeringsplasser osv).

SamspillBilUka_2	Tenk tilbake på dine møter med bilister den siste uka Forestill deg at du har møtt 100 bilister i løpet av den siste uka. Omtrent hvor mange av disse vil ha :											
◆ range:*												
	Ingen	ca 10	ca 20	ca 30	ca 40	ca 50	ca 60	ca 70	ca 80	ca 90	Alle	
	1	2	3	4	5	6	7	8	9	10	11	
latt være å overholde vikeplikten for deg som syklist i et	○	○	○	○	○	○	○	○	○	○	○	1

SamspillBilUka_2	Tenk tilbake på dine møter med bilister den siste uka Forestill deg at du har møtt 100 bilister i løpet av den siste uka. Omtrent hvor mange av disse vil ha :												
kryss													
latt være å overholde vikeplikten for deg som syklist i en rundkjøring	○	○	○	○	○	○	○	○	○	○	○	2	
oppptrådd som om han/hun tydeligvis ikke har sett deg	○	○	○	○	○	○	○	○	○	○	○	3	
plassert seg slik i vegbanen at du ikke kom forbi	○	○	○	○	○	○	○	○	○	○	○	4	
latt være å stoppe for deg da skulle sykle over et fotgjengerfelt	○	○	○	○	○	○	○	○	○	○	○	5	

BilAtferdUka	Tenk på hvordan det har vært å sykle i den siste uka I hvilken grad har du opplevd at bilister

BilAtferdUka	Tenk på hvordan det har vært å sykle i den siste uka I hvilken grad har du opplevd at bilister

◆ range:*

	1 I svært liten grad	2	3	4	5	6	7 I svært stor grad	
	1	2	3	4	5	6	7	
er oppmerksomme på deg som syklist i trafikken?	○	○	○	○	○	○	○	1
viser hensyn til deg som syklist?	○	○	○	○	○	○	○	2
overholder vikeplikten til deg som syklist?	○	○	○	○	○	○	○	3
overholder trafikkreglene	○	○	○	○	○	○	○	4

SamspillFotUka	Med tanke på hvordan du har opplevd å sykle den siste uka, og dine møter med FOTGJENGERE Forestill deg at du har møtt 100 fotgjengere i løpet av den siste uka. Omtrent hvor mange av disse vil ha :

◆ range:*

Ingen	ca 10	ca 20	ca 30	ca 40	ca 50	ca 60	ca 70	ca 80	ca 90	Alle
1	2	3	4	5	6	7	8	9	10	11

SamspillFotUka	Med tanke på hvordan du har opplevd å sykle den siste uka, og dine møter med FOTGJENGERE Forestill deg at du har møtt 100 fotgjengere i løpet av den siste uka. Omtrent hvor mange av disse vil ha :
opptrådt som om han/hun tydeligvis ikke har sett deg	○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ 1
opptrådt uforutsigbart overfor deg som syklist	○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ 2
hindret deg i å komme frem	○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ 3
plutselig gått ut i vegbanen foran deg	○ ○ ○ ○ ○ ○ ○ ○ ○ ○ ○ 4

FotAtferdUka	Tenk på slik det har vært å sykle i den siste uka I hvilken grad har du opplevd at fotgjengere							
◆ range:*								
	1 I svært liten grad	2	3	4	5	6	7 I svært stor grad	
	1	2	3	4	5	6	7	
er oppmerksomme	○	○	○	○	○	○	○	1

FotAtferdUka	Tenk på slik det har vært å sykle i den siste uka I hvilken grad har du opplevd at fotgjengere							
på deg som syklist i trafikken								
viser hensyn til deg som syklist	○	○	○	○	○	○	○	2
overholder trafikkreglene	○	○	○	○	○	○	○	3

NestenUlykkeIuka	Har du i løpet av den siste uka opplevd en nestenulykke med bil eller fotgjenger, dvs at en av dere eller begge måtte bråbremse eller svinge unna for ikke å krasje?		
◆ range:*			
	Ja 1	Nei 2	
Med bil	○	○	1
Med fotgjenger	○	○	2

MidtveisKom	Du er nå omtrent midtveis i skjemaet. Har du noen kommentarer til skjemaet så langt?
	Open

Normer	Ta stilling til følgende påstander
◆ range:*	

Normer	Ta stilling til følgende påstander							
	1 Helt enig	2	3	4	5	6	7 Helt uenig	
	1	2	3	4	5	6	7	
Jeg har en moralsk plikt til å ta hensyn til andre trafikanter uansett hvordan de oppfører seg	○	○	○	○	○	○	○	1
Det er helt i orden at syklister bryter noen trafikkregler for å komme seg fortere frem i trafikken	○	○	○	○	○	○	○	2

Utrygg	Hvor utrygg føler du deg når...								
◆ range:*									
	1 Svært trygg	2	3	4	5	6	7 Svært utrygg	Ikke aktuelt/ vet ikke	
	1	2	3	4	5	6	7	8	
... du vanligvis sykler	○	○	○	○	○	○	○	○	1
... du sykler i rundkjøring	○	○	○	○	○	○	○	○	2

Utrygg	Hvor utrygg føler du deg når...								
... du sykler på gang-/sykkelveg	○	○	○	○	○	○	○	○	3
... du sykler i sykkelfelt	○	○	○	○	○	○	○	○	4
... du sykler i vegbanen	○	○	○	○	○	○	○	○	5

ProblemOpplevelse	Har du i løpet av den sist uka opplevd noe av følgende som syklist?				
◆ range:*					
	Nei	Ja, en gang	Ja, flere ganger	Ja, mange ganger	
	1	2	3	4	
At biler har passert altfor nærme	○	○	○	○	1
At du har blitt presset av bilister	○	○	○	○	2
At bilister har kjørt på en aggressiv og truende måte	○	○	○	○	3
At bilister har kjeftet på deg	○	○	○	○	4
At bilister har tutet på deg	○	○	○	○	5
At bilister har sprutet på deg med spyleveske	○	○	○	○	6

Atferd2	Har du selv gjort noe av følgende i løpet av den siste uka?

Atferd2	Har du selv gjort noe av følgende i løpet av den siste uka?

◆ range:*

	Nei	Ja, en gang	Ja, flere ganger	Ja, mange ganger	
	1	2	3	4	
Kjeftet på en bilist	O	O	O	O	1
Vist negative gester til en bilist	O	O	O	O	2
Presset deg inn foran en bilist slik at den har måttet bremse	O	O	O	O	3

Trening	Hvor mye av syklingen din den siste uka har vært til hverdagsturer, hvor mye til fritidsturer, og hvor mye til ren trening?

◆ range:*

	Ingen turer	Noen turer	De fleste turer	Alle turer	
	1	2	3	4	
”Hverdagsturer” (til og fra jobb/skole/fritidsaktiviteter)	O	O	O	O	1
Fritidsturer i helger osv.	O	O	O	O	2
Aktiv sports- og treningssykling	O	O	O	O	3

intensitet	Noen sykler hardt og intensivt også på hverdagsturer. Andre tar det mer med ro, og vil helst unngå å bli svett. Grad av intensitet

intensitet	Noen sykler hardt og intensivt også på hverdagsturer. Andre tar det mer med ro, og vil helst unngå å bli svett. Grad av intensitet

◆ range:*

	1 Svært rolig	2	3	4	5	6	7 Svært intensivt	
	1	2	3	4	5	6	7	
Hvor intensivt sykler du på hverdagsturer (ikke helge- og treningsturer)?	○	○	○	○	○	○	○	1

sykkeltype	Når du sykler på hverdagsturer, hva slags sykkel bruker du vanligvis

◆ range:*

Terrengsykkel	○	1
Hybrid	○	2
Racer	○	3
Leid bysykkel	○	4
Klassisk sykkel	○	5
Annet	○	6

Utstyr	Hvor ofte bruker du på dine hverdagsturer?

Utstyr	Hvor ofte bruker du på dine hverdagsturer?

◆ range:*

	Alltid 1	Ofte 2	Av og til 3	Sjelden 4	Aldri 5	
Hjelm	O	O	O	O	O	1
Lys når det er mørkt	O	O	O	O	O	2
Refleksvest/sykkeljakke	O	O	O	O	O	3
Sykkelsko	O	O	O	O	O	4
Sykkelbukse	O	O	O	O	O	5
Pulsklokke	O	O	O	O	O	6
Sykkelcomputer	O	O	O	O	O	7

personlighet	Aller sist ønsker vi å vite litt om deg som person. Nedenfor finner du en rekke påstander som passer mer eller mindre godt for ulike mennesker. Kryss av i den ruten som passer best for deg slik du vanligvis er. Ikke tenk for mye på hver oppgave, men sett et kryss i ruten du umiddelbart synes stemmer best.

◆ range:*

	1 Passer helt 1	2 2	3 3	4 4	5 5	6 6	7 Passer ikke 7	
Er pratsom	O	O	O	O	O	O	O	1
Kan være kald og fjern	O	O	O	O	O	O	O	2

personlighet	Aller sist ønsker vi å vite litt om deg som person. Nedenfor finner du en rekke påstander som passer mer eller mindre godt for ulike mennesker. Kryss av i den ruten som passer best for deg slik du vanligvis er. Ikke tenk for mye på hver oppgave, men sett et kryss i ruten du umiddelbart synes stemmer best.							
Gjør en grundig jobb	○	○	○	○	○	○	○	3
Er deprimert, nedtrykt	○	○	○	○	○	○	○	4
Er original, kommer med nye ideer	○	○	○	○	○	○	○	5
Har en tendens til å være stille av seg	○	○	○	○	○	○	○	6
Er hjelpsom og uegoistisk i forhold til andre	○	○	○	○	○	○	○	7
Har en tendens til å ha lite orden på tilværelsen	○	○	○	○	○	○	○	8
Er avslappet, takler stress godt	○	○	○	○	○	○	○	9
Har livlig fantasi	○	○	○	○	○	○	○	10
Er utadvendt og sosial	○	○	○	○	○	○	○	11
Kan noen ganger være uhøflig	○	○	○	○	○	○	○	12

personlighet	Aller sist ønsker vi å vite litt om deg som person. Nedenfor finner du en rekke påstander som passer mer eller mindre godt for ulike mennesker. Kryss av i den ruten som passer best for deg slik du vanligvis er. Ikke tenk for mye på hver oppgave, men sett et kryss i ruten du umiddelbart synes stemmer best.							
Legger planer og følger dem opp	○	○	○	○	○	○	○	13
Bekymrer seg mye	○	○	○	○	○	○	○	14
Liker å spekulere, leke med ideer	○	○	○	○	○	○	○	15
Kan være sky og hemmet	○	○	○	○	○	○	○	16
Er hensynsfull og vennlig overfor de fleste mennesker	○	○	○	○	○	○	○	17
Kan være uforsiktig	○	○	○	○	○	○	○	18
Blir lett nervøs	○	○	○	○	○	○	○	19
Har få kunstneriske interesser	○	○	○	○	○	○	○	20

PanelRekrut	Du er nå ferdig med å svare på spørsmålene. Tusen takk for dine svar! Vi ønsker å kontakte deg på nytt om ca 6 uker for å stille deg noen oppfølgingsspørsmål. Denne undersøkelsen vil bare ta ca 2 minutter å besvare. Alle som deltar i oppfølgingsundersøkelsen vil være med i trekningen av et gavekort verdt 3.000 kroner. Kan vi kontakte deg på nytt om 6 uker?

◆ range:*

Ja	○	1
Nei	○	2

sluttid	Tid for avslutning av intervjuet

◆ range:*

◆ afilla:sys_timenowf c

Fylles inn automatisk ☐☐☐☐☐☐ 1

brukttid	Tid brukt på intervjuet

◆ range:*

◆ afilla:sys_elapsedtime c

Fylles inn automatisk ☐☐☐☐☐☐ 1

SisteKommentar	Har du noen sluttkommentarer til skjemaet?

Open

Appendix B – Field questionnaire

ID:intro

Information
Jeg kommer fra Transportøkonomisk institutt, og skal gjennomføre en undersøkelse om hvordan det er å ferdes i trafikken, og samspill med andre trafikanter. Det tar ca 4 minutter å svare. Hvis du svarer, er du med i trekningen av et gavekort til en verdi av 5000 kroner. Det er selvfølgelig frivillig å delta, og dine svar er helt anonyme. Er du villig til å delta?

starttid	Tid for oppstart av intervjuet
◆ range:* ◆ afilla:sys_timenowf c Fylles inn automatisk	☐☐☐☐☐☐ 1

startdato	Dato for oppstart av intervjuet
◆ range:* ◆ afilla:sys_date c Fylles inn automatisk	☐☐☐☐☐☐☐☐ 1

random	Randomvariabel
◆ range:*	

random	Randomvariabel
◆ afilla:sys_random c Fylles inn automatisk	☐☐☐☐☐☐ 1

respondent	Respondentnummer til matching
◆ range:script:mktimerandom(\starttid.a.1,\random.a.1) ◆ afilla:sys_range c Genereres automatisk	☐☐☐☐☐☐☐☐☐ 1

Intervjusted	Registrer intervjusted
◆ range:*	
Tåsen	○ 1
Ullevål sykehus	○ 2
Akersgata	○ 3
Annet	○ 4

intervju	Hvilket intervju er dette?
◆ range:*	
Sykkel	○ 1
Gående	○ 2

intervju	Hvilket intervju er dette?

ID:sykkelspm

filter:\intervju.a=1

sykkeltype	Til Intervjuer: Registrer sykkeltype og utstyr Type sykkel	
Terrengsykkel		○ 1
Hybrid		○ 2
Racer		○ 3
Leid bysykkel		○ 4
Klassisk sykkel		○ 5
Annet		○ 6

utstyr	Utstyr	
Hjelm		☐ 1
Lys		☐ 2
Refleksvest/gul jakke		☐ 3
Sykkelsko		☐ 4

utstyr	Utstyr		
Sykkelbukse		☐	5
Sykkelcomputer		☐	6

LangSykkel	Ca hvor langt har du syklet siden turen startet? I tid		
◆ range:*			
Under 5 minutter		○	1
5 til 10 minutter		○	2
10 til 15 minutter		○	3
15 til 20 minutter		○	4
20 til 30 minutter		○	5
Over 30 minutter		○	6

VantTil	Er dette en rute du er vant til å sykle?		
◆ range:*			
Ja, sykler her hver dag/ofte		○	1
Ja, sykler her av og til		○	2
Ja, har syklet her noen få ganger		○	3

VantTil	Er dette en rute du er vant til å sykle?
Nei, dette er første gang jeg sykler her	○ 4

SamspillSykBil	Tenk på din sykkeltur i dag, og tenk på dine møter med BILISTER i ulike situasjon (kryss, rundkjøringer, avkjørsler, parkeringsplasser osv). Hvor mange ganger har du opplevd

◆ range:*

	Ingen	1 gang	2 ganger	3 ganger	4 ganger	5 ganger eller flere	
	1	2	3	4	5	6	
at en bilist tydeligvis ikke har sett deg	○	○	○	○	○	○	1
at en bilist har plassert seg slik i vegbanen at du ikke kom forbi	○	○	○	○	○	○	2

SamspillSykBil_1	Tenk på din sykkeltur i dag, og tenk på dine møter med BILISTER i ulike situasjon (kryss, rundkjøringer, avkjørsler, parkeringsplasser osv). Hvor mange ganger har du opplevd

◆ range:*

	Ingen	1 gang	2 ganger	3 ganger	4 ganger	5 ganger

SamspillSykBil_1	Tenk på din sykkeltur i dag, og tenk på dine møter med BILISTER i ulike situasjon (kryss, rundkjøringer, avkjørsler, parkeringsplasser osv). Hvor mange ganger har du opplevd						
	1	2	3	4	5	6 eller flere	
at en bilist ikke har overholdt vikeplikten for deg i et kryss	○	○	○	○	○	○	1
at en bilist ikke har overholdt vikeplikten for deg i en rundkjøring	○	○	○	○	○	○	2

SykVurdering Bil	På denne turen I hvilken grad har du opplevd at BILISTER							
◆ range:*								
	1 I svært liten grad	2	3	4	5	6	7 I svært stor grad	
	1	2	3	4	5	6	7	
er oppmerksomme på deg som syklist i	○	○	○	○	○	○	○	1

SykVurdering Bil	På denne turen I hvilken grad har du opplevd at BILISTER							
trafikken?								
overholder trafikkreglene?	○	○	○	○	○	○	○	2
viser hensyn til syklister i trafikken?	○	○	○	○	○	○	○	3
er klar over at det er syklister i trafikken?	○	○	○	○	○	○	○	4

SamspillSykFot	Tenk på dine møter med FOTGJENGERE i ulike situasjoner (kryss, fortau, overgang osv). Hvor mange ganger har du på denne turen opplevd						
◆ range:*							
	Ingen	1 gang	2 ganger	3 ganger	4 ganger	5 ganger eller flere	
	1	2	3	4	5	6	
at en fotgjenger tydeligvis ikke har sett deg	○	○	○	○	○	○	1
at en fotgjenger har opptrådt uforutsigbart	○	○	○	○	○	○	2

SamspillSykFot	Tenk på dine møter med FOTGJENGERE i ulike situasjoner (kryss, fortau, overgang osv). Hvor mange ganger har du på denne turen opplevd						
overfor deg							
at en fotgjenger har hindret deg i å komme frem	○	○	○	○	○	○	3
at en fotgjenger har gått ut i vegbanen uten å se seg for	○	○	○	○	○	○	4

SykVurderingFot	1/3 av skjema På denne turen I hvilken grad har du opplevd at FOTGJENGERE							
◆ range:*								
	1 I svært liten grad	2	3	4	5	6	7 I svært stor grad	
	1	2	3	4	5	6	7	
er oppmerksomme på deg som syklist i trafikken?	○	○	○	○	○	○	○	1

SykVurdering Fot	1/3 av skjema På denne turen I hvilken grad har du opplevd at FOTGJENGERE							
overholder trafikkreglene?	○	○	○	○	○	○	○	2
viser hensyn til syklister i trafikken?	○	○	○	○	○	○	○	3
er klar over at det er syklister i trafikken?	○	○	○	○	○	○	○	4

EgenAtferdIdag	På denne sykkelturen, har du gjort noe av følgende?				
◆ range:*					
	Nei 1	Ja, en gang 2	Ja, to ganger 3	Ja, flere enn to ganger 4	
Syklet på rødt lys	○	○	○	○	1
Syklet i vegbanen selv om det var en sykkelsti/-veg ved siden av	○	○	○	○	2
Syklet på fortau	○	○	○	○	3
Stoppet for en bilist selv om det var du som	○	○	○	○	4

EgenAtferdIdag	På denne sykkelturen, har du gjort noe av følgende?				
hadde forkjørsrett					
Krysset over fotgjengerfelt uten å gå av sykkelen	O	O	O	O	5

TryggIdag	Har du følt deg utrygg noen gang når du har syklet i dag?	
◆ range:*		
Nei	O	1
Ja, en gang	O	2
Ja, to ganger	O	3
Ja, tre ganger	O	4
Ja, flere enn tre ganger	O	5

NestenUlykkeIdag	På denne turen, har du opplevd en nestenulykke med bil, dvs at du og /eller bilisten må bråbremse eller svinge unna for ikke å krasje?	
◆ range:*		
Ja	O	1
Nei	O	2

Appendix C - NbClust

Journal of Statistical Software 19

	Name of the index in **NbClust**	Optimal number of clusters
1.	"ch" (Calinski and Harabasz 1974)	Maximum value of the index
2.	"duda" (Duda and Hart 1973)	Smallest number of clusters such that index > criticalValue
3.	"pseudot2" (Duda and Hart 1973)	Smallest number of clusters such that index < criticalValue
4.	"cindex" (Hubert and Levin 1976)	Minimum value of the index
5.	"gamma" (Baker and Hubert 1975)	Maximum value of the index
6.	"beale" (Beale 1969)	Number of clusters such that critical value >= alpha
7.	"ccc" (Sarle 1983)	Maximum value of the index
8.	"ptbiserial" (Milligan 1980, 1981)	Maximum value of the index
9.	"gplus" (Rohlf 1974; Milligan 1981)	Minimum value of the index
10.	"db" (Davies and Bouldin 1979)	Minimum value of the index
11.	"frey" (Frey and Van Groenewoud 1972)	Cluster level before index value < 1.00
12.	"hartigan" (Hartigan 1975)	Maximum difference between hierarchy levels of the index
13.	"tau" (Rohlf 1974; Milligan 1981)	Maximum value of the index
14.	"ratkowsky" (Ratkowsky and Lance 1978)	Maximum value of the index
15.	"scott" (Scott and Symons 1971)	Maximum difference between hierarchy levels of the index
16.	"marriot" (Marriot 1971)	Max. value of second differences between levels of the index
17.	"ball" (Ball and Hall 1965)	Maximum difference between hierarchy levels of the index
18.	"trcovw" (Milligan and Cooper 1985)	Maximum difference between hierarchy levels of the index
19.	"tracew" (Milligan and Cooper 1985)	Max. value of second differences between levels
20.	"friedman" (Friedman and Rubin 1967)	Maximum difference between hierarchy levels of the index
21.	"mcclain" (McClain and Rao 1975)	Minimum value of the index
22.	"rubin" (Friedman and Rubin 1967)	Minimum value of second differences between levels
23.	"kl" (Krzanowski and Lai 1988)	Maximum value of the index
24.	"silhouette" (Rousseeuw 1987)	Maximum value of the index
25.	"gap" (Tibshirani *et al.* 2001)	Smallest number of clusters such that criticalValue >= 0
26.	"dindex" (Lebart *et al.* 2000)	Graphical method
27.	"dunn" (Dunn 1974)	Maximum value of the index
28.	"hubert" (Hubert and Arabie 1985)	Graphical method
29.	"sdindex" (Halkidi *et al.* 2000)	Minimum value of the index
30.	"sdbw" (Halkidi and Vazirgiannis 2001)	Minimum value of the index

Table 2: Overview of the indices implemented in the **NbClust** package.

Figure 19 Indices in the NbClust package. The table retrieved from page 19 in the article: Charrad, M., Ghazzali, N., Boiteau, V., & Niknafs, A. (2014). NbClust: an R package for determining the relevant number of clusters in a data set. *Journal of Statistical Software, 61*(6), 1-36.

Number of clusters	2	3	4	5	6	7	8
Indices	CH Cindex Silhouette Duda PseudoT2 Beale McClain	Scott TrCovW Ratkowski Ball	Friedman	KL Hartigan Marriot TraceW Rubin DB PtBiserial	CCC Dunn SDindex		SDbw

Figure 20 Results from the NbClust analysis. Indices and corresponding optimal number of clusters

Appendix D – Script from R

```
dataset=read.csv2("/Users/ragnhildbatseba/Documents/Skole/Segregated
 or integrated bicycle infrastructure/R/TilR2406.csv", header = TRUE,
 sep = ";", quote = "\"",dec = ",", fill = TRUE, na.strings = "NA")

#assigning the independent variables the callname "dataset"
 Including header, marking missing values NA. http://www.r-
 statistics.com/2013/08/k-means-clustering-from-r-in-action/

dataset <- na.omit(dataset)

#Removing missing values listwise

wss <- (nrow(dataset)-1)*sum(apply(dataset,2,var))
for (i in 2:15) wss[i] <- sum(kmeans(dataset, centers=i)$withinss)
plot(1:7, wss, type="b", xlab="Number of Clusters", ylab="Within
 groups sum of squares")

# Creating and plotting within group sum of squares for 2:8 clusters.

library(NbClust)

# Loading NbClust library

set.seed(1)

#set initial seed to 1

nc <- NbClust(dataset, min.nc=2, max.nc=8, method="kmeans")
table(nc$Best.n[1,])
barplot(table(nc$Best.n[1,]),
          xlab="Number of Clusters", ylab="Number of Criteria",
          main="Number of Clusters Chosen by 26 Criteria")
 nc$All.index

 #Test 26 of the 30 included indexes in the NbClust package for 2:8
 clusters. Need to specify the cluster method, which is K-means.
 Create barplot and show all scores.

 set.seed(1)
 two.km=kmeans(dataset, 2, nstart=25, iter.max=1000)
 two.km$size
 two.km$centers

 # Create two clusters using the K-means method. Set initial seed to
 1. Showing cluster sizes and centers. Iter.max shows the number of
```

iterations, nstart= how many random sets that should be chosen. Repeat for three and five clusters.

```
set.seed(1)
three.km=kmeans(dataset, 3, nstart=25, iter.max=1000)
two.km$size
two.km$centers

set.seed(1)
five.km=kmeans(dataset, 5, nstart=25, iter.max=1000)
two.km$size
two.km$centers

library(xlsx)

#Load library to write cluster belonging to excel file

write.xlsx(two.km$cluster, file="twoclus.xlsx")
write.xlsx(three.km$cluster, file="threeclus.xlsx")
write.xlsx(five.km$cluster, file="fiveclus.xlsx")

#Write all three solutions to excel. Added them into the same data file together with relevant variables for Anova analysis.
```